Isolation and characterization of phenylalanine ammonia-lyase and ACC oxidase genes in cassava (Manihot esculenta Crantz)

Li, Hongying

Award date:
2002

Awarding institution:
University of Bath

ISOLATION AND CHARACTERIZATION OF PHENYLALANINE AMMONIA-LYASE AND ACC OXIDASE GENES IN CASSAVA (*Manihot esculenta* Crantz)

Submitted by Hongying Li

For the degree of Doctor of Philosophy of the University of Bath

Signed 

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

First of all, I would like to give specially many thanks to my supervisor Dr. John Beeching for providing me the opportunity to study for the PhD, his patient guidance and continuous encouragement through my whole PhD work. I am very grateful for his valuable assistance and help. Without these, I could not finish my PhD.

I would also like to thank Dr Claude Fauquet at ILTAB for providing me with the opportunity to work at ILTAB and giving me continued encouragement and advice. I am very grateful to Dr. Nigel Taylor for his valuable assistance and help. Without his help, this work would not have been so successful. I also would like to thank Dr Shi-Ping Zhang and Li-Li Chen for their generous help and support in rice transformation. I would say a big thank you to the members of the 'United Nations' of ILTAB for the cheerful time, including Yadav, Pita, Dai, Anju, Thao, Terisa, Isbel, Nicole and Koffi.

At University of Bath, I would like thanks all the people in MPG lab for their help and support. Thanks also go to Paul Jones for doing a lot of sequences for me, and Michael for taking care of my transgenic plants.

I would like to thank School of Biology and Biochemistry, University of Bath, for the Studentship, and Rockefeller Foundation for the financial support during my stay in ILTAB.

Lastly I would like to thank my husband, Yuanhuai Han for his support and encouragement.

ABSTRACT

Cassava is a vitally important tropical root crop that is the staple food for over 500 million throughout the humid tropics. Within 48 hours of harvest cassava roots start to deteriorate, which renders them unpalatable and unmarketable. This post-harvest physiological deterioration (PPD) is a major constraint to developing cassava as a crop and impacts on producers, consumers and processors. While PPD resembles wound responses in other better-studied plants, the wound healing aspects are inadequate or absent. General phenylpropanoid metabolism plays an important role during wound responses, and phenylalanine ammonia-lyase (PAL) catalyses the first committed step in the core reaction. Ethylene is well known for enhancing fruit ripening and it is one of the main components in the complex signal network of wound responses. ACC oxidase catalyses the last step in the biosynthesis of the phytohormone. In this PhD thesis, genes encoding PAL and ACC oxidase in cassava were studied.

Cassava contained at least four PAL genes, two of which were isolated and characterised in this thesis. The 2677 bp of MePAL1 was sequenced. It consisted of two 381 bp and 1748 bp exons and one 592 bp intron. The 6447 bp of MePAL2 was sequenced too, which included 2740 bp of 5'flanking region, two exons of 388 bp and 1748 bp interrupted by one 711 bp intron, and 860 bp of 3' flanking region. Comparison of among four cassava PAL genes showed high similarity in the coding region, especially in the second exon, in which the nucleotide similarity was 89%-98%.

Sequence motifs, G, P A and L, associated with stress and wound gene induction in other related genes were identified in the promoter region of MePAL2. So too were two putative TATA boxes. 840 bp of the MePAL2 promoter, containing the G, P, A and L motifs, was fused to the reporter gene β-glucuronidase (GUS) to give MePAL2-840, and transformed into cassava, as were deletion mutants (MePAL400, MePAL260 and MePAL200) of the promoter in which successive motifs had been removed. While the larger constructs expressed GUS in transgenic cassava, the smaller constructs, lacking motifs did not. The MePAL840 promoter construct was studied in detail. It showed expression in the xylem parenchyma cells of the vascular tissue of all tissues examined and, in some cases, in the cells of the epidermis and bark cambium. In the storage root,

the MePAL840 promoter was up-regulated during deterioration, expression being confined to those vascular tissues in which the visual discoloration observed during PPD occurs. The effects of designed mutations in the two putative TATA boxes were studied in transient expression assays. These experiments indicated that the second TATA box, nearest the initiator ATG, was probably the genuine one.

Additionally, the PAL promoter-GUS constructs were transformed into rice. GUS expression pattern in transgenic rice plants during development was similar to that observed in cassava, demonstrating that the promoter was active in a monocotyledenous host.

ACC oxidase was also encoded by multiple genes in cassava as shown by Southern blotting. An ACC oxidase gene, MeACO2, was isolated from a genomic library. It was 6349 bp including 4928 bp of 5' flanking region and four putative exons interrupted by three introns. The deduced amino acid sequences of MeACO2 and MeACO1 (isolated from a PPD-related cDNA library by Han, 2000) was of 90% similarity and the deduced C-terminal half of two genes were of even higher similarity. The sequences around the TATA box region and intron/exon junction in MeACO2 were extremely similar to those in three well-studied ACC oxidase genes in tomato.

CONTENTS

ABBREVIATION

4CL	4 coumarate:CoA ligase
aa	amino acid
ABA	abscisic acid
ACC	1-amino-cyclopropane-1-carboxylic acid
ACMD	African cassava mosaic disease
ACMV	African cassava mosaic virus
ACO	ACC oxidase
AIP	2-aminoindan-2-phosphonic acid
AOPP	alpha-aminooxy-beta-phenylpropionic acid
BAC	bacterial artificial chromosome
BAP	6-benzylaminopurine
bp	base pair
C4L	cinnamic acid-4-hydroxylase
CaMV	cauliflower mosaic virus
CBB	cassava bacteria blight
CBN	Cassava Biotechnology Network
cDNA	complementary deoxyribonucleic acid
CHS	chalcone synthase
CIAT	Centro International de Agricultura Tropical, in Colombia
CsCMV	cassava common mosaic virus
CsVMV	cassava vein mosaic virus
DNA	deoxyribonucleic acid
dNTP	2′-deoxyribonucleoside triphosphate
EDTA	diamino ethanetetra-acetic acid, disodium salt
Esculin	6,7-dihydroxycoumaroyl-6-β-D-glucoside
FAO	Food and Agricultural Organization
FEC	Friable embryogenic callus
GCG	Genetics Computer Group
GD medium	Gresshoff and Doy, 1974
GFP	Green fluorescent protein

GUS	β-glucorunidase
HRGP	hydroxyproline-rich glycoprotein
IAA	indole acetic acid
IITA	International Institute of Tropical Agriculture, Ibadan, Nigeria
IPTG	isopropyl β-D-thiogalactopyranoside
JA	jasmonate acid
Kb	kilobase
LB	Luria-Bertani media
mRNA	messenger ribonucleic acid
MS medium	Murashige and Skoog, 1962
MUG	4-Methyl umbelliferyl beta-D-glucuronide
NAA	α-naphthalene acetic acid
NCBI	National Center for Biotechnology Information
OD	optical density
PAL	phenylalanine ammonia-lyase
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pin2	proteinase inhibitor II
PPD	post-harvest physiological deterioration
PR	pathogenesis-related
PRPs	phosphoserine aminotransferase
PVC	polyvinyl chloride
PVP	polyvinyl pyrollidone
RACE	rapid amplification of cDNA ends
RAPD	random amplified polymorphic DNA
RFLP	ramdom fragment length polymorphism
RH	relative humidity
RNase	ribonuclease
ROS	reactive oxygen species
rpm	revolution per minute
RT-PCR	reverse transcriptase PCR
SA	salicylic acid
SAICARS	5-aminoimidasole-4-N-succinocarboxyamide ribonucleotide synthetase

SAM	S-adenosyl methionine
SAR	systemic acquired resistance
Scopoletin	6-methoxy-7-hydroxycoumarin
Scopolin	6-methoxy-7-hydroxycoumaroyl-7-β-D-glucoside
SDS	sodium dodecyl sulphate
SH medium	Schenk and Hildebrandt, 1972
TMV	tobacco mosaic virus
Tris	Tris (hydroxymethyl) methylamine
UTR	untranslated region
UV	ultra-violet
VsP	vegetative storage protein
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER ONE: INTRODUCTION

1.1 Botanical description of cassava

1.1.1 Cassava, general

Cassava (*Manihot esculenta* Crantz), otherwise known as yuca, tapioca, manioc or mandioca, is a dicotyledonous perennial plant and belongs to the botanical family *Euphorbiaceae*. The cassava plant has a chromosome number 2n=36 and polyploids are not common. The size of its haploid genome is 772 mega-base pairs, which is very small in comparison with other angiosperms (Awoleye *et al.*, 1993). Like most other members of *Euphorbiaceae* family, such as rubber (*Hevea brasiliensis*) and castor bean (*Ricinus communis*), the cassava plant contains laticifers that produce latex. The cassava plant is semi-woody shrub varying from one to five meters in height depending on the cultivar and conditions of cultivation. Propagation takes place vegetatively, via lignified, mature, tough and woody stem cuttings. After planting, the new roots are established first and axillary buds sprout to form the shoot system, which takes about three to six days. Within the first and second months, roots grow rapidly but the stems grow slowly. Then rapid development of the aerial part follows for the next three months. Three to four months later, photosynthates produced by the established leaf canopy are diverted to the root system where the excess energy is converted to starch and stored in the parenchyma of greatly thickened storage roots or tuberous roots (**Figure 1.1**). Rapid growth of the tubers occurs from the sixth month and is continued to the ninth month. The storage roots are ready for harvest about eight to eighteen months after planting depending on cultivars (Onwueme, 1978). Although some people consume the leaves as a green vegetable, the crop is cultivated primarily for the high calorific value of its storage roots.

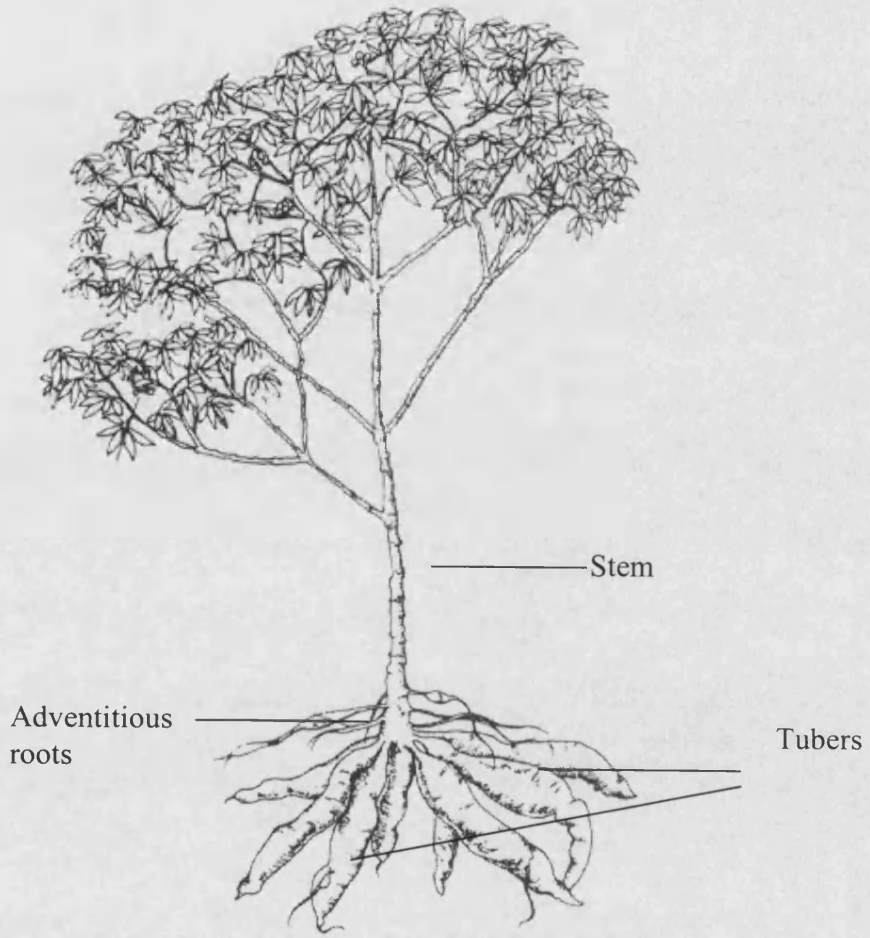


Figure 1.1 General morphology of the cassava plant (Modified from Doku, 1969)

1.1.2 Cassava origin and environmental conditions for cassava growth

Cassava is one of the oldest cultivated crops in the tropical and subtropical areas of the world (Jennings, 1976). *M. Flabellifolia* has now been confirmed as the most likely progenitor of cultivated cassava, with a centre of origin along the southern border of Amazonia (Olsen and Schaal 1999). From Latin America, cassava was introduced to Africa in the 16th century and to Asia in the late 17th century, and today it is cultivated world-wide in more than 80 countries between 30° south and 30° north of the equator. Cassava is best suited to warm (25-29°C) and humid (annual rainfall of 100-150 cm) lowland tropics, but it can be cultivated in most areas where the mean annual temperature exceeds 20°C and the annual precipitation is between 50-800 cm. It grows very poorly at temperatures below 10°C and cannot withstand frost at any time during its active growing period. It is only during the first few weeks after planting that the cassava plant is unable to tolerate drought to an appreciable extent. The best soil for cassava cultivation is a light, sandy loam soil of medium fertility, though it can grow and yield reasonably well on soils of low fertility where production of most other crops would be uneconomical. On clay or poorly drained soils, the tuber-to-shoot ratio is considerably decreased (Onwueme, 1978). Tuber formation in cassava is under photoperiodic control. Under short-day conditions, tubering occurs readily, but when the day-length is longer than 10-12 hours, tuberization is delayed and subsequent yields are low (Bolhuis, 1966; Mogilner *et al.*, 1967).

1.1.3 Root system

During the first two to three weeks of growth, adventitious roots develop at the base of the stem cutting. These roots later develop into a fibrous root system, which can penetrate 50-100 cm deep in the soil. After 30 to 60 days, only four to eight of these roots in general, begin to swell, undergo secondary thickening and then become tuberous roots (Figure 1.1). The actual number of tubers on a cassava plant and the size of a tuber depend on genotype, nutrient supply, photoperiod and temperature. A mature cassava tuber, which is generally fattest at the proximal end and tapers gently toward the distal portion, ranges in length from 15-100 cm and in weight from 0.5-2.0 kg. The cross section of a young tuber is divisible into three regions, peel, starchy flesh and

central vascular strand (**Figure 1.2**). The outermost layer, the periderm, sclerenchyma, phloem and cork cambium compose the cassava peel. Cassava peel comprises 10-20% of the tuber. The flesh, which is the central portion and consists largely of storage parenchyma, is the main storage part of cassava. This consists mostly of parenchymatous cells containing large amounts of stored starch and xylem vessels that ramify randomly throughout the flesh. Large xylem vessels and fiber tissues run through the centre of the flesh and compose the vascular strand. The edible fleshy portion makes up 80-90% of the tuber and is composed of about 62% water, 35% carbohydrate, 0.5-2% protein, 0.3% fat, 1-2% fiber and 1% mineral matter (Purseglow, 1968; Yeoh and Truong, 1996). Most of the carbohydrate is starch, which makes up 20-25% of the tuber flesh (Purseglow, 1968). The protein in cassava tuber is very low in both quantity and quality. It is relatively rich in glutamine, aspartic acid and arginine but low in methionine, lysine, tryptophan, phenylalanine and tyrosine (Oyenuga, 1968; Yeoh and Truong, 1996). Among the minerals in the tuber, phosphorus and iron predominate, with minimal amounts of calcium.

1.1.4 The leaves, stems and flowers of cassava

The shoot system of cassava develops from axillary buds located on the nodes of a cutting. The cassava stem can grow up to five meters but dwarf varieties may by only one meter tall. The stems vary considerably in colour and maybe whitish, brown or dark brown, and are usually woody with very large pith. In older parts of the stem, prominent knob-like leaf scars are present, marking the nodal position where leaves were originally attached. The distance between nodes varies with cultivar and also with environmental conditions. The cross sections of the cassava stem correspond to that of a typical dicot stem. In the primary state, before the onset of secondary growth, it is surrounded by an epidermis, beneath which is the cortex. Internal to the cortex, the vascular bundles are arranged in a ring. Each bundle consists of phloem and xylem separated by a thin layer of vascular cambium. As secondary growth progresses, the stem increases in girth. Lenticels and cork develop on the surface, and a large amount of xylem is formed.

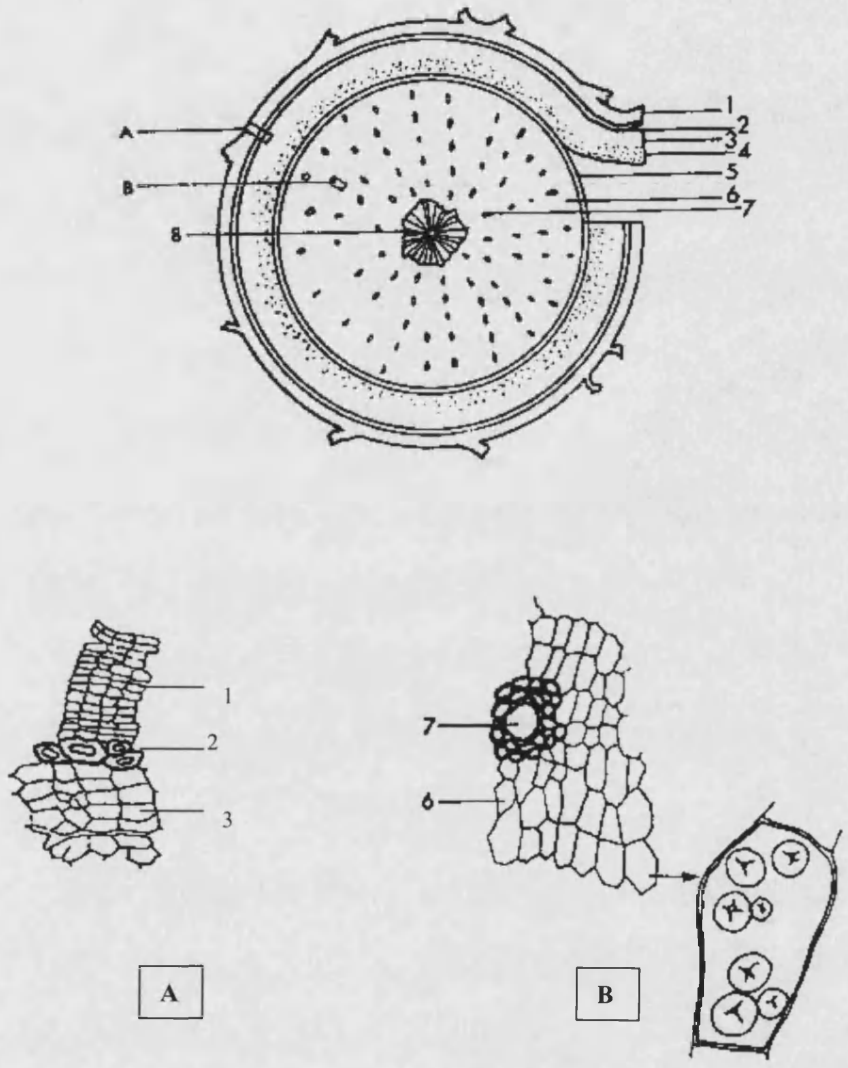


Figure 1.2 Transverse section of a young cassava tuber (modified from Doku, 1969)

A:enlarged 1,2 and 3. B: enlarged 7 and 8 with large starch grains

1: Periderm 2: Sclerenchyma 3: Cortical parenchyma 4: Phloem (1-4:peel) 5: Cambium 6: Storage parenchyma 7: Xylem vessel 8: Xylem vessel and fibres

The leaves of cassava are spirally arranged on raised nodal portions on the stem. The phyllotaxis is 2/5. Each leaf is subtended by 3-5 stipules, each stipule being about 1 cm long. The leaves of cassava are deciduous and under the best conditions, each leaf exists for a few months before it falls off. Consequently, the lower portions of the cassava stem are devoid of leaves. Cassava leaves contain a high level of protein, about 29.3-32.4% dry weight (Awoyinka *et al.*, 1995), and vitamins, especially Vitamins A and C. The leaves are therefore much richer than the tuber in proteins and vitamins so, especially in Africa, people consume the leaves as a green vegetable. The leaf petiole varies in length from 5 to 30 cm dependent on cultivar and position of the leaf. In general the petiole is longer than lamina.

Flowering in cassava is frequent and regular in some cultivars, but in other cultivars, flowering is rare or non-existent. Cassava is monoecious; both male and female flowers occur on the same plant. In each inflorescence, the female flowers open first, while the male flowers do not open until about a week later. In nature and in the process of plant breeding, propagation by seed is quite common, though in agricultural production it is propagated exclusively from stem cuttings.

1.2 Agronomic characteristics and economic roles of cassava

Cassava is the most important staple food crop in the tropics and the fourth most important calorie source after rice, sugarcane and maize. Although relatively unknown in developed countries, cassava is widely cultivated in over 80 countries of the topics and subtropics in Africa, Asia and South America (CIAT, 1992) and provides basic staple food for over 500 million people daily from a total of 16 million hectares. The world annual cassava production increased from 70 million tonnes in 1960 to 167.7 million tonnes in 1999 (Table 1.1) (FAO, 2000). FAO forecasts that production will rise to nearly 210 million tonnes by 2005. Cassava production has risen steadily during the past decades, mainly due to increased growing area and introduction of new high-yielding varieties. (FAO, 2000). The largest cassava producers are Nigeria, Thailand, Indonesia, Brazil and Democratic Republic of Congo (Table 1.1).

Table 1.1 development of cassava production during 1970-1999, million tones (FAO, 2000)

	1970	1980	1990	1999
World	**98.7**	**124.1**	**152.4**	**166.7**
Africa	**40.5**	**48.3**	**64.1**	**85.5**
Congo Dem.Rep.	10.3	13.1	17.0	15.0
Nigeria	10.2	11.5	17.6	30.4
Asia	**23.2**	**485.8**	**52.0**	**50.2**
Indonesia	10.6	13.6	16.3	15.4
Thailand	3.4	16.5	21.9	20.3
Latin America	**34.7**	**29.2**	**33.7**	**30.8**
Brazil	29.5	23.5	25.4	22.5
Colombia	1.2	2.2	4.0	2.0

The photosynthetic system in cassava is very efficient, which enables it to accumulate as much as 250 x 10^3 cal/ha/day; 25% and 40% greater than in rice and maize, respectively (Splittstoesser, 1992). The cassava tuber is one of the most efficient producers of starch, with 85% of the dry-matter of the storage root being starch. Most of the harvested tubers are used for human food, either fresh or in various processed forms. The rest is processed to animal feed and industrial products, such as meal, flours, chips (or pellets), starch alcohol, glucose and others.

There are also many useful agronomic characteristics of the cassava plant, which account for its widespread cultivation. The cassava plant can tolerate poor soils and drought, and it is less affected by pests and diseases compared to other crops. It allows acceptable harvests even on marginal and eroded soils, which are unable to support any other crop. The tubers can be harvested at any time from 8 to 24 months after planting, which is an important safeguard against unexpected food shortage or famine, though old roots become a little more fibrous. These features make cassava one of the most productive crop plants and the cheapest known source of starch. When grown on fertile soils under high radiation, cassava yields are comparable to rice, maize and sugarcane in terms of dry weight.

1.3 Constraints in cassava production

1.3.1 Diseases and pests

Despite its relative tolerance to diseases and pests, cassava is susceptible to a number of biotic and abiotic constraints. Various pests and diseases are estimated to cause 20%-50% yield losses world-wide, and locally they can lead to total crop failures. The common cassava diseases include cassava mosaic virus, bacteria blight, bacterial stem rot, brown leaf spots, white thread and others. Cassava bacterial blight (caused by the bacterium *Xanthomonas axonopodis* pv. *manihotis*) is one of the main biotic constraints in cassava cultivation world-wide and heavy infestations of bacterial blight can destroy the whole crop (Lozano, 1979; Boher and Verdier, 1994; Mahungu *et al.*, 1994). It is recognized by the presence of watersoaked, angular spots, blight and wilting of the leaves, vascular necrosis in the stem and root, and the die-back of the shoots. It is usually spread through infected cuttings. By breeding, some resistant varieties have been produced, but so far the resistance obtained appears to be effective only under low infestation pressure (Cooper *et al.*, 1995). The bacterium, *Xanthomonas axonopodis* pv. *manihotis* penetrates the host through stomata, or through small surface wounds. Once in the host it destroys the spongy mesophyll of the leaf and enters the vascular tissues. It then spreads systemically to the vascular tissues in all parts of the plant. Most of this movement occurs through the xylem vessels.

Two major virus diseases are cassava common mosaic virus (CsCMV) and African cassava mosaic disease (ACMD), transmitted by white flies (*Bemisia tabaci*), infected tools, or stem cuttings. They cause losses of up to 40-50% of total yields throughout the whole world and can locally destroy the whole cassava (Thresh *et al.*, 1994; Otim-Nape, 1995). Several pests such as mealybugs (*Phanacoccus manihoti*), white flies (*Bemisia*), stem borer (*Chilomina Clarkei*), green mites, red spider and root-knot nematodes (*Meloidogyne incognita*) affect cassava yield, and sometimes cause severe damage. Green mites can cause up to 80% losses and nematodes even can cause up to 98% yield losses (Yaninek, 1994).

1.3.2 Cyanogenesis

Another important problem is the cyanogenic nature of cassava. All parts of the cassava plant produced hydrogen cyanide when damaged. Hydrogen cyanide is the breakdown product of cyanogenic glycosides, linamarin and lotaustralin, which are synthesized from valine and isoleucine in leaf tissue and transported to roots. Mechanical damage of the cassava tissue can deglycosylate linamarin by linamarase to acetone cyanohydrin, which can break down to acetone and HCN spontaneously or by hydroxynitrile lyase activity. The cyanogenic potential of cassava is highly variable between individual plants, and even between individual roots of one plant, and it is also influenced by the environment (Bokanga, 1994; Bokanga *et al.*, 1994). Cassava leaves contain the highest cyanogenic glycoside level of all cassava tissues, which is 5.0 g linamarin/kg fresh weight. The level of total root linamarin ranges from 100 to 500 mg/kg fresh weight depending on the cultivars (McMahon *et al.*, 1995). Cassava cultivars are divided into two groups, bitter types and non-bitter types (or sweet types or cool types), according to the taste of the roots. King and Bradbury (1995) discovered that cyanogenic glycosides were the sole contributor of bitterness in the parenchyma but isopropyl-β-D-glucopyranoside contributed more to the bitterness of cortex in two out of six cultivars studied. The presence of cyanogenic glycosides in cassava has been a problem for the consumption of cassava (Jones, 1998). Various disorders in tropical areas were found to be linked with the consumption of cassava (White *et al.*, 1998), food-poisoning or even death can be caused by consumption of unprocessed or not fully-processed bitter cassava. Cassava roots are usually detoxified during processing by baking or drying or retting (steeping the roots in water for several days).

1.3.3 Post-harvest physiological deterioration

Rapid post-harvest physiological deterioration (PPD) is another major constraint to cassava production. The cassava tuber, unlike other tropical tuber crops such as yam, potato, sweet potato, cannot be stored for more than a few days after harvesting as the tubers begin to deteriorate rapidly (even within 24 hours), which renders them unpalatable and unmarketable. On average, deterioration occurs within 2-3 days after harvest due to physiological processes, and is followed by microbial deterioration

within 5-7 days (secondary deterioration) (Plumbley and Rickard 1981). The symptoms of PPD are initially blue/black vascular streaking that develops from wound sites and along xylem strands followed by discoloration of the storage parenchyma. Microscopic observation revealed that the initial response to wounding in cassava roots was occlusion of the xylem vessels and production of phenolic compounds in the storage parenchyma. The main components of the occlusions are carbohydrates, lipids and lignin-like material (Rickard *et al.*, 1981; Rickard, 1983).

A number of methods have been used to prolong the storage life of cassava roots. Cassava roots can be harvested on a small scale only when they are needed, and immediately consumed, processed or marketed to avoid rapid post-harvest deterioration. However, if the roots are left under the ground for too long they become more lignified and fibrous, leading to low palatability (Onwueme, 1978; Ravi *et al.*, 1996). Dipping in water of 60°C for 45 min is effective in inactivating PPD or inhibiting the occurrence of vascular streaking during the 5-day observation after the treatment (Averre, 1967). Cassava roots can be stored for several weeks at low temperatures of 0-5°C (Montaldo, 1973; Booth, 1976). Cassava roots treated at high temperature (35°C) and high humidity (80-85% relative humidity) for two to four weeks after harvest, a process called 'curing', can be stored up to four more weeks (Booth, 1976). Fresh harvested roots with minimal mechanical wounding can also be stored for up to 2 to 4 weeks in an acceptable condition for marketing in boxes packed with moist sawdust (Booth, 1977). Cutting off the top of the plant at 20-30 cm above the stem base two to three weeks prior to harvesting, termed pruning, can reduce susceptibility to PPD (Lozano *et al.*, 1978; Tanaka *et al.*, 1984). Roots treated with fungicide (thiabendazole) and packed within 3 hours after harvest, can be stored for 15 days in polythene bags at about 30°C (Wheatley *et al.*, 1989). Commercially, roots can be dipped in paraffin wax at 90 to 95°C for short periods such as 45 seconds to prolong the shelf life up to 1 to 2 months (Ravi *et al.*, 1996). However, the storage methods mentioned above are neither convenient, nor economical, nor suitable for marketing and processing on a large scale, and the deterioration problem remains to be solved to enable farmers to make full use of their land to improve their livelihoods.

Physiological and biochemical changes during the storage of cassava roots were investigated in the early studies of the phenomenon. Cassava roots are inevitably wounded when they are taken off the plants during harvest and usually suffer further wounding such as abrasions during harvesting and transporting. It was observed by Booth (1976) that PPD was initiated from wound sites and spread through the whole root, and there was a link between the degree of the mechanical damage of the roots and the level of PPD. Another obvious change during storage was water loss of the roots. The water loss rate from roots in low humidity was more than two fold of that in high humidity (Marriott *et al.*, 1978). Reducing water loss of the roots by storing the roots in polyethylene bag or at high humidity (85% RH) can significantly reduce PPD (Wheatley, 1989; Rickard, 1985). Oxygen was identified to be one of the important factors involved in the development of PPD. PPD was significantly reduced when roots were stored in the absence of oxygen, such as in pure CO_2 and propane (Marriott *et al.*, 1978). An increase in respiration rate was observed one day after harvest in low humidity but there was no change in respiration rate at high humidity (Marriott *et al.*, 1978). It was also shown by Data *et al* (1984) that the respiration rate in the roots increased two fold within one day after harvest. Ethylene production increased during PPD but exogenous application of ethylene did not affect the development of PPD (Plumbley *et al.*, 1981; Hirose *et al.*, 1984).

The blue-brownish discoloration, the first visible symptom of PPD, is a well-known phenomenon associated with the oxidation of phenolic compounds upon wounding in other plant systems such as potato tubers. Therefore, early studies on PPD were concentrated on changes in the profiles of phenolic compounds and enzymes related to phenolic metabolism.

An increase in coumarins was detected at an early stage of PPD, before the occurrence of discoloration. These coumarins included scopoletin (6-methoxy-7-hydroxycoumarin), scopolin (6-methoxy-7-hydroxycoumaroyl-7-β-D-glucoside), esculin (6,7-dihydroxycoumaroyl-6-β-D-glucoside) and two conjugates containing scopoletin and esculin respectively (Rickard, 1981; Tanaka, 1983). Intense fluorescence, mainly due to these coumarins, was observed in the roots before the visible symptoms of PPD. Total phenol content expressed as chlorogenic acid, total

flavanols expressed as catechin and leucoanthocyanidins expressed as cyanadin chloride, increased during the development of physiological deterioration (Rickard, 1985). The time course studies of scopoletin, scopolin and esculin during storage of cassava roots showed that scopoletin was produced first and peaked along with scopolin at 24 hours; esculin peaked at 40 hours after wounding (Wheatley and Schwabe, 1985). In order to understand the link between phenolic compounds and PPD, Wheatley and Schwabe (1985) tested a number of phenolic compounds for their effect on the development of PPD. These compounds included cinnamic acid, *p*-coumaric acid, caffeic acid, ferulic acid, coumarin, umbelliferone, esculetin, arbutin, catechol and scopoletin. They showed that only scopoletin caused a significant deterioration after 18 hours incubation, leading to visual symptoms of deterioration similar to those of the naturally occurring deterioration. More interestingly, they found that roots from pruned plants responded to scopoletin vigorously, whereas the roots stored in low or no oxygen condition did not respond. Wheatley and Schwabe thus suggested that the absence of oxygen during storage might result in the loss of a scopoletin precursor or inactivation of related enzymes, whereas pruning was effective possibly due to reduction in scopoletin or absence of some factors involved in the primary oxidation.

Enzymes involved in the biosynthesis of phenolic compounds and their oxidation were investigated. The key enzyme of phenylpropanoid biosynthesis, phenylalanine ammonia-lyase (PAL), showed increased activity peaked at about 72 hours (± 36) after harvest (Richard, 1982; 1985; Tanaka *et al.,* 1983). PAL activity was significantly lower in roots from pruned plants than those from un-pruned plants one to two days after harvest (Tanaka *et al.,* 1984), which may explain the lower content of scopoletin in roots from pruned plants as Wheatley and Schwabe (1985) discovered. The activities of peroxidase and polyphenol oxidase were also increased during PPD, which was consistent with the appearance of discoloration in the vascular tissues (Rickard, 1985).

Although still poorly understood, the metabolic changes observed during PPD resemble those observed during normal plant wounding and defense responses in other plant systems. There have been extensive studies on wounding and defense responses in other plant systems, by reference to which the PPD in cassava may be analyzed.

1.4 Plant defense responses upon wounding and pathogen attack

Plant defense against wounding or pathogen attack involves several different layers of activity or response. The waxy cuticle, together with the periderm beneath the cuticle provides physical barriers against bacterial and fungal pathogens. Strengthened cell walls by increased suberin, lignin, and cell wall proteins such as peroxidase and hydroxyproline-rich glycoproteins (HRGPs) and their in-solublization are associated with plant defense responses (Bach and Seitz 1997; Faulkner and Kimmins 1975; Grosskopf *et al.*, 1991 Showalter *et al.*, 1985). Lignification and suberisation associated with callose deposition were effective mechanisms that reinforced host barriers in the phloem in cassava defense against *Xanthomonas axonopodis* pv. *manihotis* (Kpemoua *et al.* 1996). Occlusion of xylem vessels by polysaccharide gels and tyloses can prevent the spread of vascular pathogens (Cooper, 1981). Secretion of phenol-like molecules by tyloses were only observed in a cassava variety resistant to *X. axonopodis* pv. *manihotis* (Kpemoua *et al.* 1996). Some low molecular weight anti-microbial compounds synthesized before pathogen attack provide plants with another layer of defense, such as saponins (a group of triterpenes that disrupt fungal membranes by binding to sterols) and simple phenylpropanoids (Taiz and Zeiger, 1989). Upon pathogen attack, a series of responses are induced such as production of reactive oxygen species, synthesis of pathogenesis-related proteins (PR proteins) and phytoalexins, reinforcement of cell walls by synthesizing lignin or callose and cross-linking cell wall proteins. Salicylic acid has been shown to be a secondary signal required for expression of both localized hypersensitive response and systemic acquired resistance, which activate the expression of those defense genes (Ryals *et al.*, 1994, Delaney *et al.*, 1994, Mur *et al.*, 1996). Reactive oxygen species may lead to lipid peroxidation, enzyme inactivation and degradation of nucleic acids, contributing cell death as part of hypersensitive response or killing the pathogen directly (Lamb and Dixon, 1997). Hydrolytic enzymes such as glucanases and chitinases are typical PR proteins, which attack fungal cell walls. Phytoalexins include a diverse group of secondary metabolites with strong antimicrobial activity synthesized and accumulated upon pathogen attack. Proline-rich proteins such as HRGPs in cell walls become oxidatively cross-linked by an H_2O_2-mediated reaction upon pathogen attack (Bradley *et al.*, 1992).

Plants respond to wounding in a similar way as to pathogens, such as the generation of reactive oxidative species, insolubilization of proline-rich cell wall proteins, lignification and suberization of cell walls to reinforce or repair the extracellular matrix at the wound sites (Bowles, 1990; Bradley *et al.*, 1992; Dixon and Lamb, 1990; Orozco-Cardenas and Ryan, 1999). A number of wound-responsive (WR) genes have been isolated and characterized, such as proteinase inhibitors, JR (jasmonate acid-responsive) genes, prosystemin and VSP (vegetative storage protein) (Farmer and Ryan, 1990; Peña-Cortes *et al.*, 1988; Leon *et al.*, 1998; McGurl *et al.* 1992; Mason and Mullet, 1990). Several possible chemical inducers of wound response have been identified. Oligosaccharides released from fungal or plant cell walls during plant-pathogen interactions can activate plant defense response at very low concentrations (John *et al.*, 1997). It was suggested recently that oligosaccharides might also function as a wounding signal based on the characterization of a wound- and systemin-inducible polygalacturonase in tomato leaves (Bergey *et al.*, 1999). In tomato leaves, a short polypeptide (18-aa) called systemin signals the expression of wound-responsive genes (McGurl *et al.*, 1992). Systemin is processed from the C terminus of a 200-aa precursor prosystemin (McGurl *et al.*, 1992), is more than one million-fold more active than oligosaccharides and is rapidly transported through the plant (Pearce *et al.*, 1991). The influence of wound signals including oligosaccharides, systemin, plant hormone jasmonic acid (JA), abscisic acid (ABA) and ethylene, and electrical signals, on the regulation of WR genes have revealed a complex signaling network during wounding response. Two regulatory systems for wound responses are identified, one in solanaceous plants and another in *Arabidopsis thaliana*. The solanaceous model suggested that during the wound response, oligosaccharides and systemin activate JA synthesis, activating WR gene expression, ethylene amplifies the JA signal and results in a maximal response (O'Donnell *et al.*, 1996). A JA-independent wound signal pathway seems to exist in tomato (O'Donnell *et al.*, 1998) and was confirmed in *Arabidopsis* (Titarenko *et al.*, 1997). In the *Arabidopsis* model (**Figure 1.3**) proposed by Rojo *et al.* (1999), oligosaccharides are released from plant cell walls upon wounding, which block JA synthesis, and activate ethylene production and JA-independent genes. The ethylene suppresses JA-induced gene expression in the

wounded leaves but relatively short distance movement of oligosaccharides and ethylene diffusion may fail to suppress JA-induced gene expression in systemic leaves.

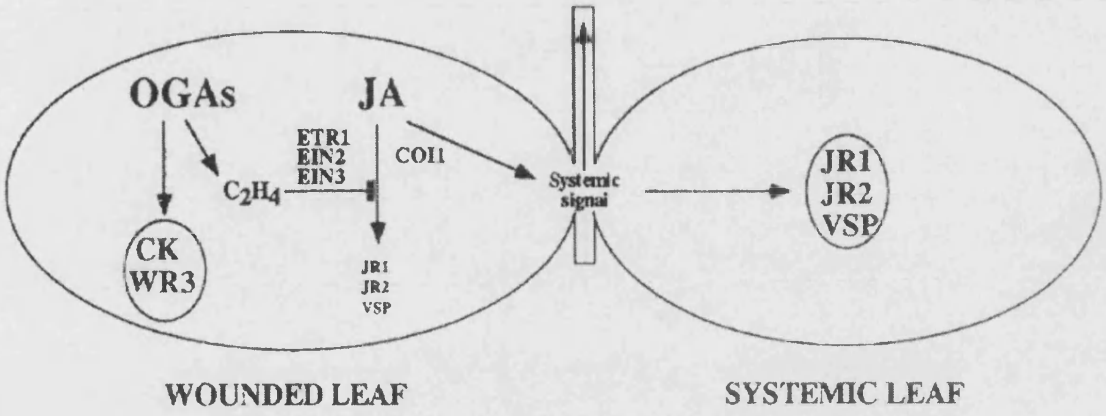


Figure 1.3 A model for wound signal transduction in *Arabidopsis thaliana* (Rojo *et al.*, 1999)

A putative role of endogenous oligosaccharide (OGA), jasmonic acid (JA) and ethylene (C_2H_4) in gene activation in the damaged (wounded leaf) and non-damaged (systemic leaf) leaves of wounded *Arabidopsis* plants is depicted. The components of the ethylene and JA transduction pathways (ETR1, EIN2, EIN3, and Col1) are located at positions likely for their action. Arrows indicate positive action; blunted lines indicate negative regulation. The set of gene active in the respective tissues is encircled.

As described above, ethylene is one of the main components in the complex signal network of wounding response. In many of these responses upon pathogen attack and wounding, general phenylpropanoid metabolism plays a central role. This PhD project was, therefore, focused on the key genes involved in these two aspects.

1.5 Phenylpropanoid metabolism and phenylalanine ammonia lyase

1.5.1 Phenylpropanoid pathway

A large and diverse array of organic compounds in plants seems to have no direct function in growth and development. These compounds are known as secondary metabolites to distinguish them from the primary metabolites such as chlorophyll, amino acids, nucleotides, simple carbohydrates or membrane lipids, which are involved directly in growth and development. Secondary metabolites were thought to be

functionless for many years, until recently many secondary metabolites were suggested to have important roles in plant defense against herbivores and pathogens. The secondary metabolites are divided into three classes according to their distinct chemical features, which are terpenes (lipids synthesized from acetyl CoA or basic intermediates of glycolysis), phenolics and nitrogen-containing compounds (such as alkaloids) synthesized from amino acids.

Most abundant classes of secondary phenolic compounds are derived from the phenylpropanoid pathway. All phenylpropanoids were derived from cinnamic acid, which is formed from phenylalanine by its deamination to form *trans*-cinnamic acid, catalyzed by phenylalanine ammonia-lyase (PAL). The cinnamic acid is converted to *para*—coumaric acid by addition of a hydroxyl group at a *para* position on the aromatic ring. Coumaric acid and derivatives such as ferulic acid and sinapic acid can be catalyzed by 4-coumarate:coenzyme A ligase (4CL) to produce cinnamoyl-CoA. The series of reactions from deamination of phenylalanine to the synthesis of cinnamoyl-CoA, involving the enzymes PAL, cinnamic acid-4-hydroxylase (C4H) and 4CL, are considered as a common core reactions (**Figure 1.4**). From this core reaction, it may lead to different branch pathways such as the synthesis of lignin, suberin, flavonoids, isoflavonoids, coumarins, esters and salicylic acid.

Phenolic compounds play a great diversity of roles in plants. The release of simple phenolics may affect the growth of neighbouring plants. Compounds such as caffeic acid and ferulic acid have been found to accumulate in soil in appreciable amounts and have been shown to inhibit the germination and growth of many plants (Rice, 1987). A wide range of phenylpropanoids has been identified as involved in defense and stress responses such as UV light, pathogen attack, signaling and wounding (**Figure 1.5**). Flavonoid derivatives, which function as pigments, provide plant flowers and fruits with colours including red, pink, purple and blue to attract animals for pollination and seed dispersal (Taiz and Zeiger 1998). Flavonoids may also protect plants against UV light and function as phytoalexins. *Arabidopsis* mutants lacking chalcone synthase activity and producing no flavonoids were more sensitive to UV radiation than the wild type (Li *et al.*, 1993). Many simple phenolics have important roles in plants as defenses against herbivores and pathogens. Some phenolics such as furanocoumarins are

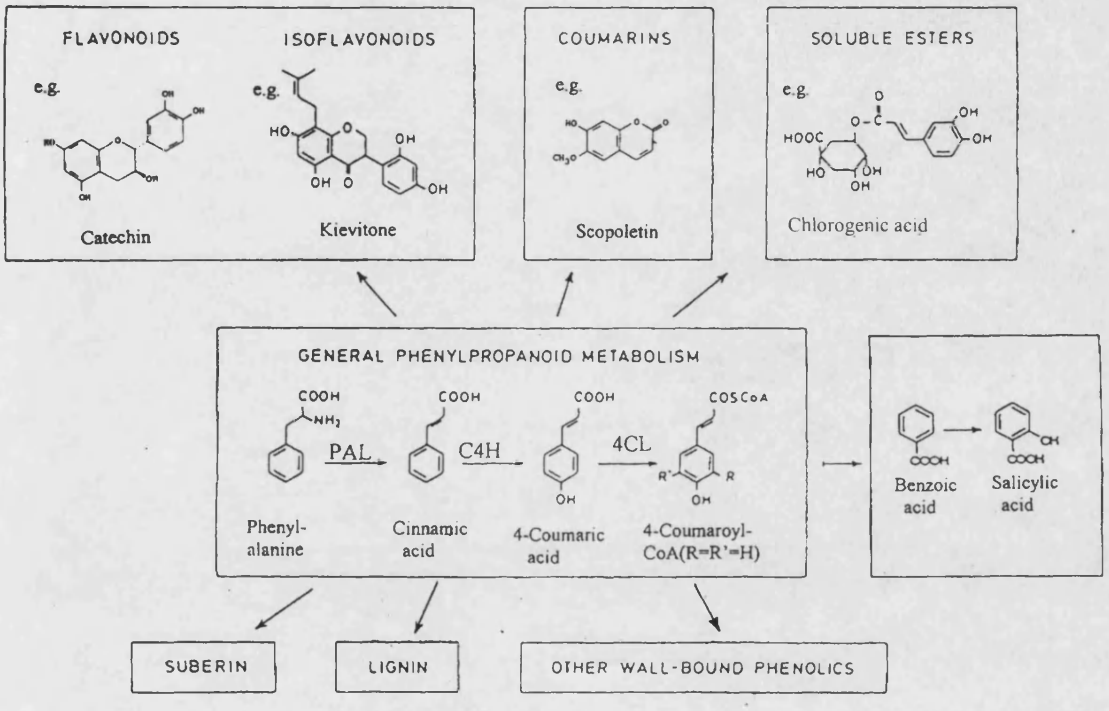


Figure. 1.4 General phenylpropanoid metabolism and major branch pathways.

PAL: phenylalanine ammonia-lyase

C4H: cinnamic acid-4-hydroxylase

4CL: 4-coumarate:coenzyme a ligase

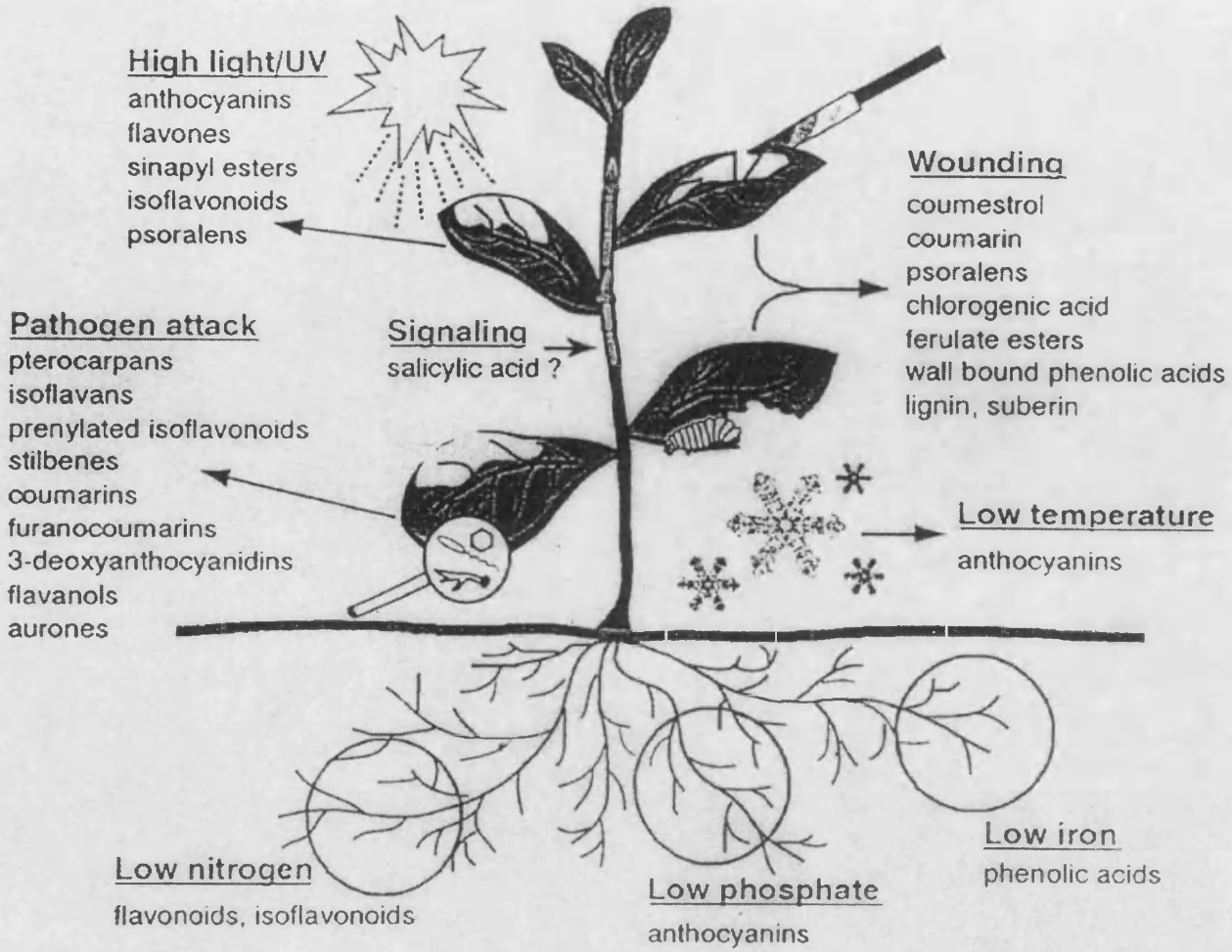


Figure 1.5 Examples of stress- induced phenylpropanoids (Dixon and Paiva, 1995)

activated by UV light, exhibiting phototoxicity to insect herbivores (Sandberg and Berenbaum, 1989). Many phenylpropanoids can be increased greatly to a high level around the infection site to concentrate toxic compounds in response to pathogen attack, such as pterocarpans, isoflavans, prenylated isoflavonoids, stilbenes, psoralens, coumarins, flavonols and aurones (Bennet and Wallsgrove, 1994; Dixon *et al.*, 1995). From cinnamic acid, benzoic acid can be synthesized. One of benzoic acid derivatives, salicylic acid, is believed to be part of a signalling process that results in systemic acquired resistance (SAR) to plant pathogens (Mur *et al.*, 1997).

Many phenylpropanoid compounds are induced in response to wounding or to feeding by herbivores. Increased levels of coumestrol and coumarin are toxic to potential herbivores (Smith, 1982). Wounding induced cholorogenic acid, alkyl ferulate esters, and cell-wall-bound phenolic esters may act directly as defense compounds or may serve as precursors for the synthesis of lignin, suberin, and other wound-induced polyphenolic barriers (Hahlbrock and Scheel, 1989). Lignin is the most abundant organic substance after cellulose in plants, it is a highly branched polymer of phenylpropanoid alcohols including coniferyl, coumaryl and sinapyl alcohols. Lignin is deposited in the walls of specific cells such as the tracheids and vessel elements of the xylem. Lignin strengthens stems and vascular tissue, and allows water and minerals to be conducted through the xylem without collapse of the tissue. Lignification in response to infection or wounding could form a physical barrier to block the pathogen or seal the wound (Boudet *et al.*, 1995). Suberin is a matrix of hydroxy fatty acids polymerized along with phenolics primarily (poly)hydroxycinnamates (Bernards and Lewis, 1998). It is deposited in the Casparian strip of the root endodermis, forming a barrier between the apoplast of the cortex and the stele. It is found principally in periderm tissues of both nonwoody plants, such as sweet potato and woody plants, such as oak (David and Lewis, 1991). Suberization of the cells at a wound site would form a physical barrier to reduce water loss and enhance defense against pathogens.

1.5.2 The role of phenylalanine ammonia-lyase in phenylpropanoid metabolism

Classical biochemical and molecular techniques have been used to isolate and purify many of the enzymes and their genes involved in phenylpropaniod biosynthesis to

analyse their function. Phenylalanine ammonia-lyase (PAL, E.C. 4.1.3.5.), which catalyses the deamination of L-phenylalanine to form *trans*-cinnamate, is one of most studied enzymes and genes of phenylpropaniod metabolism as it is the first and key enzyme of the pathway and controls the flux of primary metabolites into this major secondary metabolic pathway. It was first discovered in barley seedlings by Koukol and Conn in 1961 and since then numerous papers related to this enzyme has been published and it has become the most studied enzyme concerned with secondary metabolism in plants. This enzyme has been demonstrated to be present in all higher plants tested as well as in some yeast species (Hanson and Havir, 1981, Camm and Towers 1973, cited by Johns, 1984; Hahlbrock and Scheel, 1989), but not in animals. The subcellular location of PAL is mainly cytoplasmic, though it may also be loosely associated with the membrane-bound cinnamate hydroxylase or benzoate synthase systems (Hanson and Havir, 1981; Amrhein and Zenk, 1971, as cited by Jones, 1984).

The regulatory role of PAL in phenylpropanoid metabolism is shown by the fluctuation of some phenolic compounds associated with changes in PAL in response to stresses such as pathogen attack. In the systemic resistance to *Blumeria graminis f. sp. tritici* (Bgt) induced by benzo (1,2,3) thiadiazole-7-carbothioic acid S-methyl ester (BTH) in the susceptible wheat cv. Monopol, the activities of the enzyme phenylalanine ammonia-lyase (PAL) and peroxidase and cell wall-bound phenolic compounds (i.e. coumaric and ferulic acids) were higher than in untreated leaves (Stadnik and Buchenauer, 2000). Inhibition of PAL by AOPP (alpha-aminooxy-beta-phenylpropionic acid) reduced the localized accumulation of autofluorogenic compounds and suppressed the resistance induced by BTH. Therefore, it was suggested that a rapid accumulation of phenolic compounds associated with PAL activities at the attempted Bgt penetration sites might be involved in the resistance induced by BTH. When rice leaves were sprayed with *Pseudomonas fluorescens*, a substantial increase in phenylalanine ammonia-lyase activity was observed one day after treatment and the phenolic content of rice leaves also increased to a maximum at 4 days after *P. fluorescens* treatment (Meena *et al.,* 2000). Gamma-irradiation of Moroccan Citrus fruits (*Citrus clementina* Hort. ex. Tanaka), which can reduce the losses due to peel injury 'pitting', enhanced the synthesis of total phenolic compounds and accumulation

of flavonoids and *p*-coumaric acid and was correlated with phenylalanine ammonia-lyase activity (PAL) during storage. (Hafida *et al.*, 2000).

Extractable PAL activity as well as the total amount of phenolics in soybean (*Glycine max.* (L.) Merr) roots increased 24 h after plant transfer to low temperature. High pressure liquid chromatography analyses revealed the presence of six phenolic acids in soybean roots: p-hydroxybenzoic, vanillic, syringic, anisic, *p*-coumaric and ferulic (Janas *et al.*, 2000). The regulatory role of PAL in phenolic metabolism is also demonstrated by its effect on accumulation of lignin and suberization. Specific inhibition of PAL activity by 2-aminoindan-2-phosphonic acid (AIP) led to reduction of lignin associated with secondary wall thicking in isolated *Zinnia* mesophyll cells (Nakashima *et al.*, 1997). AIP-inhibited PAL activity in radish seedlings altered lignin composition, increased the syringyl to guaiacyl (S/G) ratio in cotyledons and decreased the ratio in hypocotyls plus roots (Chen and McClure, 2000). In tomato, rapid suberin coating in the xylem in plants resistant to wilt fungus *Verticillium albo-atrum* and significantly less vascular coating in susceptible plants upon the infection were highly correlated with PAL activity (Lee *et al.*, 1992). Inhibition of PAL activity by S-carvone, affected wound healing in potato tuber, presumably by inhibiting suberization, but when S-carvone was removed PAL activity increased and subsequently suberized cell layers were observed (Oosterhaven, 1995).

The accumulation of salicylic acid (SA) is a requirement for systemic acquired resistance (SAR), as demonstrated by the inability of tobacco plants transformed with a bacterial gene encoding salicylate hydroxylase (*nahG*) to express SAR (Gaffney *et al.*, 1993). It was shown that accumulation of salicylic acid and 4-hydroxybenzoic acid in phloem fluids of cucumber during systemic acquired resistance to *Pseudomonas syringae py. syringae* was preceded by a transient increase in PAL activity in petioles and stems (Smith-Becker *et al.*, 1998). This also indicated that PAL can affect the downstream pathways branches from the core reactions.

The role of PAL in phenylpropanoid biosynthesis was well demonstrated in transgenic tobacco plants containing a bean PAL2 gene. Over-expression of PAL in transgenic tobacco increased levels of the hydroxycinnamic acid ester chlorogenic acid but not of

the flavonoid rutin, Howles *et al* (1996) then suggested that PAL was the key control point for flux into chlorogenic acid. Suppression of PAL expression in transgenic tobacco plants led to much lower level of free SA compared to control plants (not containing the bean PAL2 transgene) and these plants failed to develop systemic acquired resistance in response to infection by tobacco mosaic virus (Pallas *et al.*, 1996), which demonstrated the role of PAL in the production of derivatives from phenylpropanoids.

1.5.3 Phenylalanine ammonia-lyase genes and their expression

The genes encoding PAL have been extensively studied over the past two decades in order to understand the spatial and temporal expression of PAL genes. PAL has been isolated and sequenced from more then 30 plant species, including a gymnosperm, pine, monocots, and herbaceous and woody dicots, including *Arabidopsis* (Mauch-Mani and Slusarenko, 1996, Ohl *et al.,* 1990), French bean (Cramer *et al.*, 1989, Sablowski *et al.,* 1995, Shufflebottom *et al.,* 1993), parsley (Lois *et al.,* 1989), pea (Yamada *et al.,* 1992), poplar (Subramaniam *et al.,* 1993), rice (Minami *et al.,* 1998, Zhou *et al.*, 1995), sweet potato (Tanaka *et al.,* 1989) tobacco (Fukasawa-Akada *et al.,* 1996) and tomato (Lee *et al.,* 1992). In most plant species, PAL is encoded by a multigene family with 2 to 6 members (Lois *et al.*, 1989; Cramer *et al.,* 1989; Frank and Vodkin, 1991; Gowri *et al.,* 1991; Lee, 1992; Lois *et al.,* 1992; Minami, 1989; Minami *et al.,* 1993; Hedrick *et al.,* 1990; Subramaniam *et al.,* 1993; Tanaka *et al.,* 1989; Wanner *et al.,* 1995) except that there were over 40 PAL genes in potato (Joos and Hahlbrock, 1992) and only one in loblolly pine (Whetten and Sederoff, 1992). The structures of most cloned plant PAL genes are similar; they have a single intron at a conserved position, and a long highly conserved second exon. One exception to this is PAL3 of *Arabidopsis*, which contained two introns (Wanner, *et al.*, 1995).

PAL genes are expressed at multiple times and places during plant development, and in response to environmental stimuli (Rickey, 1991; Lois, 1992; Collinge and Slusarenko, 1987) and other various stresses (Joos and Hahlbrock, 1992; Edward *et al.,* 1985; Ryder *et al.,* 1990). These expression patterns are generally correlated with sites of accumulation of phenylpropanoid products. As described above, general

phenylpropanoid reactions or core reactions can lead to a number of branch pathways, and, as the first enzyme in the core reactions, PAL has been shown to play important roles in controlling some of the branch pathways. Therefore, it is not difficult to predict the location and timing of PAL gene expression and activity. For example, PAL mRNA accumulation in the vascular systems of plant organs is associated with the deposition of lignin in xylem tracheary elements, expression in petals is associated with the biosynthesis of anthocyanin pigments. Pathogen- activated expression is associated with the biosynthesis of salicylic acid and an array of other phenolic compounds at infection sites. Numerous PAL genes have been isolated and characterized over the past two decades, which adds to the understanding of the spatial and temporal expression of PAL genes.

1.5.3.1 Developmentally regulated PAL gene expression in other plant systems

The transcripts of individual PAL genes show relatively different patterns of accumulation (Liang 1989; Lois 1989). In general, PAL mRNA is typically abundant in roots and stems or petioles, low in leaves and lowest in more mature leaves (Joos *et al.*, 1992; Liang *et al.*, 1989; Subramaniam *et al.*, 1993; Yamada *et al.*, 1992).

PAL transcripts detected by RNase protection in bean (*Phaseolus vulgaris* L.) showed that though all three genes were expressed at high levels in roots, only PAL1 and PAL2 were expressed at medium levels in shoots and only PAL1 was expressed at low levels in leaves, while PAL3 was not expressed in the leaves and shoots and PAL2 was not expressed in leaves at all. Strikingly, PAL2 was expressed at very high levels in petals, where PAL1 was only very weakly expressed and PAL3 was not expressed (Liang *et al.*, 1989). Transcripts of both PSPAL1 and PSPAL2 from pea (*Pisum sativum*) accumulated to high levels in roots and at moderate levels in stems but to very low levels in the upper part of the plant that included leaves and flower organs (Yamada, 1992). PAL transcript levels were much higher in flowers and roots than in leaves and stems of mature tobacco (Fukasawa-Akada, *et al.*, 1996). PAL transcripts accumulate differentially during flower and leaf maturation, in that mRNA levels decrease during flower maturation but increase during leaf maturation. In leaves, PAL mRNAs accumulated rapidly after wounding. In poplar trees, the highest expression of PAL was observed in young stems, apical buds and young leaves. Expression was lower in older

stems and undetectable in mature leaves. Cellular localization of PAL expression by *in situ* hybridization showed very high levels of expression in sub-epidermal cells of leaves early during leaf development. In stems and petioles, expression was associated with sub-epidermal cells and vascular tissues (Subramaniam *et al.*, 1993). In rice, transcripts of the PAL gene, ZB8, accumulated to a high level in stems, moderately high in roots and weak in leaves (Zhu *et al.*, 1995).

Although it is common for members of a gene family to show unique expression patterns, the expression patterns of PAL1 and PAL2 from *Arabidopsis thaliana* were both qualitatively and quantitatively very similar in different plant organs and under various inductive conditions (Shufflebottom, *et al.*, 1995). Both mRNAs were most abundant in roots and stems, moderately abundant in flowers, and least abundant in leaves.

1.5.3.2 PAL gene expression in response to biotic and abiotic stresses

PAL was one of the first identified plant 'defense genes' and was found to be induced by biotic and abiotic stresses. The expression of PAL genes was induced in suspension-cultured cells by elicitors from fungal cell wall and in plants by pathogen attack. PAL expression was strongly activated during the hypersensitive response in tobacco, potato, parsley and rice infected by tobacco mosaic virus (TMV), *Phytophthora infestans* and *Phytophthora megasperma* f.sp. *glycinea*, respectively (Cuypers *et al.*, 1988; Pellegrini *et al.*,1994; Schulze-Lefert *et al.*,1989).

Northern blot analyses showed that PAL gene expression was induced in tobacco leaves upon TMV infection or salicylic acid treatment (Kang *et al.*, 1998). The accumulation of both PSPAL1 and PSPAL2 transcripts in pea was induced in epicotyl tissues by treatment with the fungal elicitor isolated from a pea pathogen, *Mycosphaerella pinodes* (Yamada, 1992)

All three PAL genes (PAL1, 2 and 3) in French bean were induced by mechanical wounding of hypocotyls, though the levels of PAL1 and PAL2 transcripts were three and six fold that of PAL3 (Liang *et al.*, 1989). Fungal infection activated only PAL1

and PAL3. Illumination of etiolated hypocotyls activates PAL1 and PAL2, but not PAL3.

Wounding of leaf tissues strongly induced the transcriptional activity of the rice PAL, ZB8 gene (Zhu *et al.*, 1995). As early as one hour after treatment, the ZB8 transcripts started to increase and reached the peak level about 24 hour after wounding. For elicitor treatment, the time for maximum induction of PAL transcripts was about 3 hours.

Infection of mature potato leaves with the pathogenic fungus *Phytophthora infestans*, resulted in a dramatic and transient induction of PAL mRNA (Joos and Hahlbrock, 1992). The relative timing of PAL-1 and PAL-2 mRNA expression, differed in compatible (fungus virulent, plant susceptible) but not in incompatible interactions (fungus avirulent, plant resistant). Wounding of leaves caused an extremely rapid and transient induction of both PAL mRNA species (Joos and Hahlbrock, 1992). Sharp expression peaks of three PAL genes were detected two hours after wounding in potato leaves, which returned to original levels about six hours later (Joos and Hahlbrock, 1992). The *de novo* synthesis of PAL mRNA was rapidly induced by wounding in potato tubers (Rickey and Belknap, 1991).

UV-containing white light, fungal elicitor, and wounding all strongly induced the transient accumulation of PAL mRNAs in parsley (Logemann, *et al.*, 1995). The induction behavior and the timing of changes in the PAL1, PAL2 and PAL3 mRNA amounts were almost identical. PAL4 mRNA showed different expression patterns from the other three PAL genes. Upon wounding, PAL4 mRNA was strongly induced in roots but not in leaves, and it showed a short-life response to elicitor treatment.

1.5.3.3 PAL genes in cassava

In cassava, the partial sequence of a PAL gene, MePAL, was amplified and cloned from genomic DNA of cultivar Mcol 22 using degenerate primers designed from conserved region of PAL genes from other plants (Pereira *et al.*, 1999). The PCR clone was 520 bp in length and had homology to the second exon (middle of coding region) of PAL genes from other plants such as pea, alfalfa and soybean. Southern

hybridisation of genomic DNA from different cassava cultivars using a MePAL probe, showed different band patterns among the cultivars, especially between cultivars from South America and Africa/Asia, indicating a divergence among cultivars. The MePAL gene was expressed in cassava leaves, stems, petioles, storage roots but not in roots. It was expressed at higher levels in young tissues compared to mature tissues. Mechanical wounding of cassava leaves and roots lead to an accumulation of MePAL mRNA. MePAL was induced by mechanical wounding in leaves and the highest level of MePAL mRNA was observed between 12 to 18 hours after wounding. An increased mRNA level was observed 4 hours to 12 hours after inoculation with *Xanthomonas cassavae.*

Another two cassava PAL cDNA clones, MePAL1 and MePAL3, were isolated from a cDNA library (using bean PAL2 as a probe), which was constructed from cassava storage roots undergoing physiological deterioration (Han, 2000). The sequence of MePAL1 and MePAL3 were 2253 bp and 1126 bp (partial) respectively and share 79% similarity in nucleotide sequence and 93% similarity in amino acid sequences. These two PAL genes shared high identity in deduced amino acid sequences with PAL genes in other plant species, especially dicotyledonous tree plants such as poplar and lemon. Using reverse-transcription-PCR (RT-PCR), it was shown that MePAL1 was expressed in cassava young leaves, stems and fibrous roots but it was not detected in unwounded or healthy storage tubers. However, MePAL1 was expressed in tubers 8 hours after harvest.

Although three PAL genes have been isolated in cassava, their expression patterns were not fully explored. MePAL consisted of only part of the gene sequence, which is highly conserved not only in amino acid sequence but also in nucleotide sequence between different PAL genes in the same family. Therefore, Northern hybridisation may reveal the total accumulation of all the similar PAL genes in cassava, rather than the expression pattern of MePAL itself. To obtain gene specific expression pattern, Han (2000) used competitive RT-PCR with MePAL1-specific primers. However, it was difficult to carry out spatial expression pattern of PAL genes mainly due to the nature of cassava. The cassava plant contains laticifers that produce latex, which make expression analysis such as tissue-printing difficult. Although it was possible to analyse

gene expression in the tuber by tissue-printing, it was not possible to locate the expression to cells. Reporter genes fused to promoters have proved to be a very useful tool for studies of spatial expression patterns of genes. Therefore, in the project it is proposed to isolate a PAL promoter and to analyse the properties of the promoter and its expression pattern by promoter-GUS fusion.

1.5.4 Properties of phenylalanine ammonia-lyase promoters in other plants

A promoter, as a control region of a gene, is the critical component of gene expression and regulation. They have been extensively studied among genes from *E.coli* to higher plants and mammals. As an important gene with roles in development, defense and stress response, the functional properties of PAL promoter have been widely studied in transgenic plants using a PAL promoter fused with reporter gene such as GUS or green fluorescent protein gene (GFP). These promoters include those from the dicot plants bean, parsley, tobacco, *Arabidopsis,* poplar tree, and the monocot rice (Bevan *et al.,* 1989; Gray-Mitsumune, *et al.,* 1999 Kawamata *et al.,*1997; Liang *et al.,*1989; Mauch-Mani *et al.,* 1996; Ohl *et al.,* 1990; Shufflebottom *et al.,* 1993; Zhu *et al.,* 1995). Most of these PAL promoters were analysed in heterologous hosts, mainly in tobacco. Only a few PAL promoters were studied in both homologous and heterologous hosts to compare the expression patterns (Gray-Mitsumune, *et al.,* 1999). The focus of promoter analysis has been on the identification of potential regulatory regions (*cis*-acting sequences). This was normally achieved by constructing a deletion series of promoter fragments, which were then fused with a reporter gene whose level of expression is then analysed *in vivo*.

As expected, the reporter gene fused to PAL promoters has been shown to be developmentally regulated and stress-activated in transgenic plants in developing xylem, pigmented portions of petals, epidermal and sub-epidermal cell layers, root tips, and in response to wounding and pathogen infection. Different spatial and temporal patterns of reporter gene expression and induction by environmental stress in transgenic plants highly reflected the level of the transcripts of these PAL genes. Therefore, these expression patterns were also generally correlated with sites of accumulation of phenylpropanoid products.

The most well-studied PAL promoter in plants is the bean PAL2 promoter. The functional properties of the promoter have been well characterized by analysing GUS activity and GUS expression pattern in transgenic tobacco, *Arabidopsis* and potato plants (Bevan *et al.*, 1989; Liang *et al.*, 1989; Shufflebottom *et al.*, 1993). Bevan *et al* (1989) and Liang *et al* (1989) first reported the GUS expression driven by 1.1 kb bean PAL2 promoter in transgenic tobacco, which showed that GUS activity in these transgenic plants was very similar to that of endogenous PAL2 transcripts in bean, with very high levels in petals, roots, shoots and very low levels in ovaries and leaves. GUS activity was very high in the root and shoot apical meristems and in xylem differentiation, indicating active PAL2 promoter in these tissues. GUS activity also suggested that the promoter activity was enhanced in leaf, stem and root upon wounding and light induction. The promoter was then analysed in detail for *cis*-acting elements and domains by promoter deletions fused with GUS (Leyva *et al.*, 1992; Hatton *et al.*, 1995). Leyva *et al* (1992) showed that a region from –289 to –74 contained *cis*-elements for xylem expression and a negative element that suppressed the activity of upstream elements located between -480 to –289 required for expression in phloem and perivascular parenchyma. Deletion of the xylem elements did not affect the tissue specificity of the promoter in petals or expression in leaf primordia and nodes. It was also shown clearly that 5' deletions of the promoter reduced the activity of the promoter but did not affect expression patterns. Using 5' deletions and short internal deletions fused with GUS, Hatton *et al* (1995) defined functional domains in the PAL2 promoter and proposed that three AC-elements, together with a G-box, might interact to direct the complex tissue-specific expression patterns of the PAL gene. Footprinting and mobility shift assay studies of the minimal promoter revealed potential *cis* elements which coincided with the functional domains defined by small deletions and promoter fusions. Meanwhile the bean PAL3 promoter was analysed to compare with the PAL2 promoter in expression pattern and *cis*-elements (Hatton *et al.*, 1996). Bean PAL2 and PAL3 promoters conferred expression in overlapping sets of tissue types in transgenic tobacco. The PAL3 promoter contained motifs that resemble two AC *cis*-elements, which were required for tissue-specific expression of the PAL2 promoter. Analysis of mutated PAL3 promoter-GUS constructs in transgenic tobacco suggested that the

functions of these motifs in the PAL3 promoter were also necessary for tissue-specific expression of the PAL3 promoter.

Another well-studied plant PAL promoter was PSPAL1, which is one of the three genes in the pea (*Pisum sativum*) PAL family (Kawamata *et al.*, 1997; Murakami *et al.*, 1997). Tissue-specific expression of the promoter was investigated in transgenic tobacco carrying a PAL promoter fused with the GUS gene. GUS activity was very high in tobacco roots, particularly in meristematic tissues and the pigmented region of petals, but very low in mature leaves. GUS was expressed at wound sites in leaves. A rapid induction of GUS expression was also observed along with the development of an hypersensitive cell death area when leaves were inoculated with nonpathogens, *Phytophthora capsici, P. boehmeriae* and *Erisiphe graminis* f. sp. *hordei*. In contrast, inoculation of pathogen *P. nicotiana* showed slow formation of GUS expression zone along with the lesion development.

The ZB8 PAL promoter from rice, the first to be analysed among monocot PAL promoters, was investigated in transgenic rice and tobacco to understand its regulation by development and various environmental cues, such as UV, wounding and fungal elicitor (Zhu *et al.*, 1995). High levels of GUS activity were observed in stems, moderate levels in roots and low levels in leaves of transgenic rice and tobacco plants. Histochemical analysis indicated that in transgenic rice the promoter was active in root apical tips, lateral root initiation sites, and vascular and epidermal tissues of stems and roots. Consistent with the expression patterns of endogenous PAL genes, wounding of rice and tobacco leaf tissues, and infection of tobacco leaves with tobacco mosaic virus (TMV), enhanced GUS activity. GUS activity was strongly induced by fungal cell wall elicitors in transgenic rice suspension-cultured cells. These observations demonstrated that the regulation mechanisms of PAL gene expression during plant development and in response to environmental stresses were conserved between monocots and dicots (Zhu *et al.*, 1995).

In trees, the activities of PAL1 and PAL2 promoters from poplar have been studied in transgenic poplar and tobacco (Gray-Mitsumune *et al.*, 1999). The PAL1 and PAL2 promoters were 75% identical over about 600 bp of the sequenced regions, and each

contained two copies of AC-rich putative *cis*-acting elements. The PAL2 promoter showed the highest activity (indicated by GUS activity) of expression in roots and young leaves and stems in transgenic poplar, low in woody stems, and was weak or absent in developing secondary xylem. High GUS activity was also detected in epidermal or sub-epidermal cells as well as in primary xylem and phloem fibres in young leaves and stems. The patterns of PAL2-GUS expression reflected the mRNA accumulation patterns of PAL1 and PAL2. In transgenic tobacco, PAL1-GUS and PAL2-GUS fusions directed similar patterns of expression in developing primary xylem of leaves, stems, and other organs, and in secondary xylem of stems. Contrary to PAL1 and PAL2 expression patterns in poplar, no expression of either fusion (PAL1-GUS and PAL2-GUS) was observed in epidermal or sub-epidermal cell layers, meaning that PAL2-GUS expression in tobacco does not accurately reflect all aspects of PAL2 expression in poplar. This is the only reported case so far that GUS expression patterns of PAL promoters in homologous host plants were different from those in heterologous hosts.

Using a GUS fusion, it was shown that the PAL1 promoter from *Arabidopsis* was activated early in seedling development and in adult plants, where it was strongly expressed in the vascular tissues of roots and leaves, but it was not active in the root tip or the shoot apical meristem (Ohl, *et al.,* 1990). In flowers, GUS expression was located in sepals, anthers, and carpels, but not in petals. GUS transcripts from the PAL1-GUS gene fusion were induced by wounding, $HgCl_2$-stress, and light. Analysis of the regulatory properties of 5' deleted promoters showed that the proximal region of the promoter to -290 was sufficient for the full tissue-specific pattern of expression and that the proximal region to -540 was responsive to environmental stimuli. Negative and positive elements were located between -1816 and -823 and between -823 and -290, respectively.

The promoter sequences of all the PAL gene family members were analysed in parsley, which is the only example so far among PAL gene families in plants (Logemann *et al.,* 1995). It was shown that around the TATA box and transcription start site (positions 133 to +20), the four PAL genes were remarkably similar in nucleotide sequence. In this region PAL1, 2 and 3 were just a few base pairs different. However, in regions

further upstream and downstream the four genes differed greatly, except for PAL1 and PAL2, which were almost identical throughout the regions analysed. Upstream from the TATA boxes, the only significant sequence similarities shared by all four parsley PAL gene promoters were P, A and L boxes, which had been identified by *in vivo* DNA footprinting (Hahlbrock *et al.*, 1995). Sequences of P, A and L box in parsley were CTCCAACAACCCC, CCGTCC and TCTCACCTACC, which were also called AC rich elements or AC elements in other plants PAL promoters. These elements are also present in the TATA-proximal promoter regions (though at various relative locations) of two parsley genes of 4 coumarate:CoA-ligase (4CL), another phenylpropanoid enzyme like PAL. It was found that none of these elements alone, or the promoter region containing all of them together, conferred elicitor or light responsiveness on a reporter gene in transient expression assays. The elements appeared to be necessary but not sufficient for elicitor- or light-mediated PAL gene activation, similar to the situation reported for 4CL in parsley (Logemann *et al.*, 1995).

In summary, PAL promoters are active in vascular tissues and can be induced or enhanced by biotic and abiotic stresses. Some *cis*-acting elements in the promoters are conserved and the arrangement of the P, A and L elements seems to affect the expression patterns.

1.6 ACC Oxidase

1.6.1 Biological effects of ethylene

The functions of ethylene in plants have been extensively studied and it has been established that ethylene is involved in many aspects of plant growth, development, senescence and responses to various environmental stimuli.

Ethylene can break dormancy and stimulate the germination of dormant seeds (Esashi, 1991). Ethylene is involved in regulation of the growth and development of shoot and root. Exogenous ethylene caused the so-called 'triple response' in seedlings, reduction in root and hypocotyl elongation, swelling of the hypocotyl and abnormal directional growth (Neljubov, 1901 cited by Reid, 1995). Epinastic growth of petioles is another

typical symptom of plants in response to either exogenous or endogenous ethylene (Bradford and Yang, 1980).

Ethylene is well known for enhancing fruit ripening. Ethylene treatment, or the use of inhibitors of ethylene biosynthesis such as aminoethoxyvinylglycine (AVG), or inhibitors of its action such as 2,5-norbornadiene (NBD), and silver ions have been used to control fruit ripening (Yang, 1987; Theologis, 1992). It was first reported by Grierson *et al.* (1986) that ethylene was involved in regulating the expression of several ripening-related genes in tomato. In 1990, ethylene production during fruit ripening was reduced and fruit ripening was delayed by down-regulating the ACC oxidase gene in tomato by introducing an antisense ACC oxidase gene into the plant (Hamilton *et al.*, 1990). Since then, manipulation of fruit ripening has been approached with the genetic engineering of ethylene production (Grierson and Schuch, 1993).

Large quantities of ethylene were produced during the senescence of leaves and flowers (Smart, 1994; Stead, 1992). A rapid senescence occurred in broccoli (*Brassica oleracea L.*) floral tissues after harvest, which was associated with an increase in ethylene production by florets (Pogson *et al.*, 1995). Leaves of the ethylene-resistant mutant *et alr* displayed delayed senescence processes (Zacarias and Reid, 1990). Exogenous application of inhibitors of ethylene action can delay senescence, such as the extension of vase life of cut flowers by silver thiosulfate (Reid *et al.*, 1980). Sense and antisense ACO transgenes introduced into *Torenia fournieri* L. extended flower longevity. The primary transformants showed longevity of 2.5 to 7.1 days, whereas the longevity of the wild type was 2.0 days (Aida *et al.*, 1998).

Increased ethylene production is one of the wound responses in plants (Yang and Hoffman, 1984). Mechanical wounding of iceberg lettuce (*Lactuca sativa* L.) by cuts or punctures caused an increase in ethylene production (Ke and Saltveit, 1989). The activities of ACC synthase and ACC oxidase have been shown to be enhanced by mechanical wounding in a number of species (Bleecker *et al.*, 1988; Liang *et al.*, 1992; Liu *et al.*, 1997; B. Blume, unpublished results).

Similarly, a rapid accumulation of ethylene also occurs during plant defense against pathogens, and a number of wound-related and / or defence-related genes such as cell wall protein hydroxyproline-rich glycoprotein (HRGP), antifungal hydrolases chitinase and β-1,3-glucanase have been shown to be regulated by ethylene (reviewed by Boller, 1990). An ethylene responsive element AGCCGCC was identified in the promoter of PRB-1b gene (basic-type pathogenesis-related protein) and PR-5 (osmotin-like protein) gene in tobacco. Mutation of the sequence inhibited the induction of GUS expression driven by these promoters (Sessa *et al.,* 1995; Sato *et al.,* 1996).

1.6.2 Ethylene biosynthesis and ACC oxidase

Ethylene is produced via the following pathway: Methionine→S-adenosyl methionine (SAM)→ 1-aminocyclopropane-1-carboxylic acid (ACC)→Ethylene (**Figure 1.6**). The conversion of SAM to ACC is catalysed by ACC synthase (S-adenosyl-L-methionine methythioadenosine-lyase, EC4.44.1.14) and the oxidation of ACC to ethylene is catalysed by ACC oxidase (EC 1.4.3). These two enzymes and the corresponding gene families have been intensively studied.

The first ACC oxidase cDNA clone was identified through the function of a ripening-related cDNA clone TOM13 in tomato. TOM13, isolated from a ripening-related cDNA library made from tomato fruit (Slater *et al.,* 1985), was expressed in ripe fruit, wounded green fruit and wounded leaves, but not in green fruit and unwounded leaves (Smith *et al.,* 1986). The function of TOM13 was initially identified by expressing the antisense fragment of the cDNA using CaMV 35 S promoter in transgenic tomato plants, which decreased ethylene biosynthesis and ACC oxidase activity in ripening fruit and wounded leaves (Hamilton *et al.,* 1990). The function of the gene was confirmed by expressing the cDNA in yeast (Hamilton *et al.,* 1991), which enabled the transgenic yeast to convert ACC to ethylene.

Since the isolation of ACC oxidase genes in tomato, ACC oxidase genes have been isolated from a number of plant species including avocado (McGarvey *et al.,* 1990), carnation (Wang and Woodson, 1991), petunia (Tang *et al.,* 1993), peach (Callahan *et al.,* 1992), mustard (Wen *et al.,* 1993), apple (Ross *et al.,* 1992), melon (Balague *et al.,*

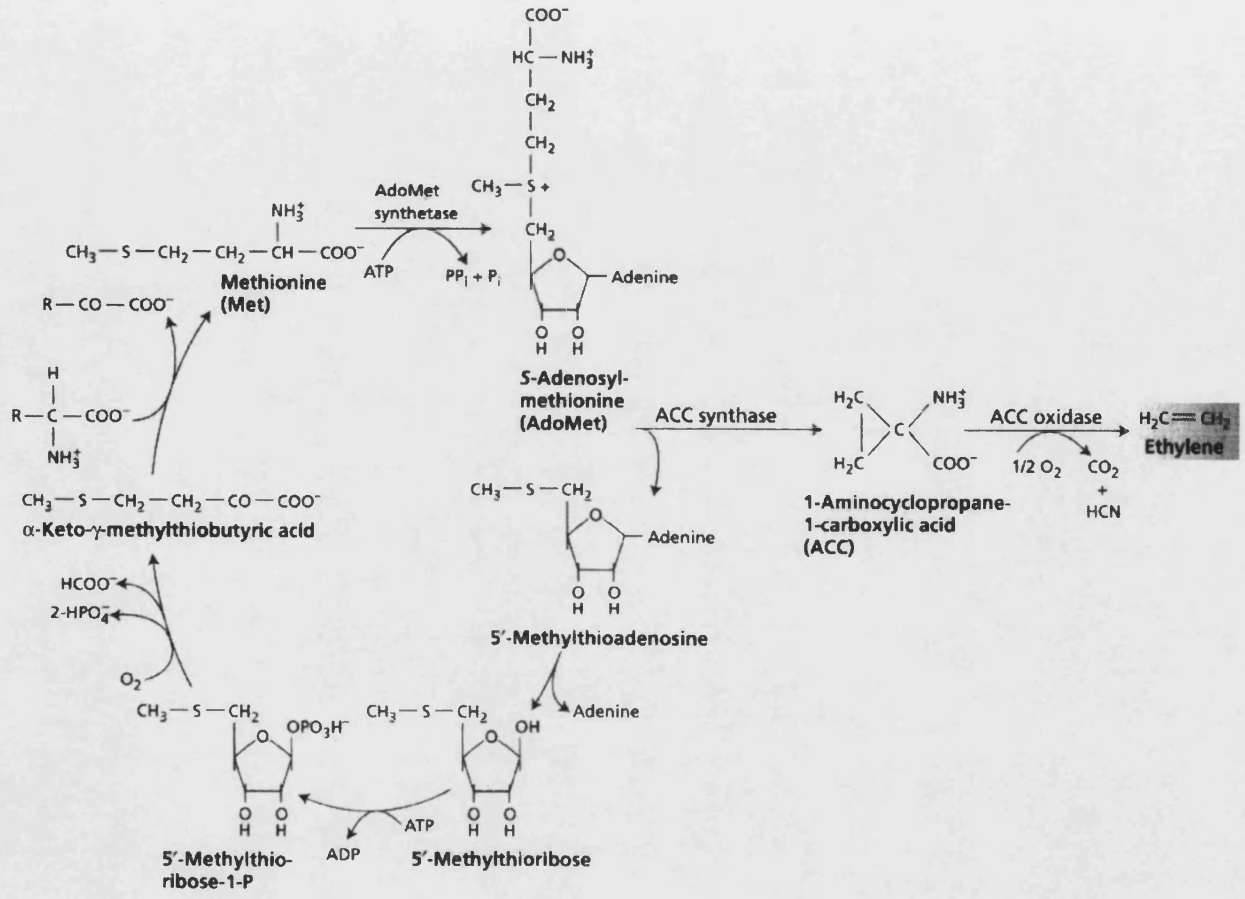


Figure 1.6 Ethylene biosynthetic pathway (modified from Taiz and Zeiger, 1998). The amino acid methionine is the precursor of ethylene. The last step in the pathway, the conversion of ACC to ethylene, requires oxygen and is catalyzed by the enzyme ACC oxidase

1993), *Arabidopsis* (Gomez-Lim *et al.,* 1993), kiwifruit (MacDiarmid and Gardner, 1993), orchid (Nadeau *et al.,* 1993), pea (Peck *et al.,* 1993), mung bean (Kim and Yang, 1994), *Pelargonium* (Wang *et al.,* 1994) and broccoli (Pogson *et al.,* 1995), banana (Liu and Pua, 1997), sunflower (Liu *et al.,* 1997), pineapple (Cazzonelli *et al.,* 1998), cucumber (Shiomi *et al.,* 1998), apricot (Mbegui-A-Mbeguie *et al.,* 1999).

Most of the clones isolated from fruits showed increased expression at the onset of fruit ripening (Callahan *et al.,* 1992; Ross *et al.,* 1992; Balague *et al.,* 1993; Barry *et al.,* 1998; Cazzonelli *et al.,* 1998; Mbegui-A-Mbeguie *et al.,* 1999). In line with the increased evolution of ethylene upon mechanical wounding, increased expression of ACC oxidase gene was observed in a number of species such as tomato, peach and broccoli during wound response (Holdsworth *et al.,* 1987; Callahan *et al.,* 1992; Pogson *et al.,* 1995).

The importance of ethylene in the wound response, senescence and prolonging shelf-life in other plants or crops suggests its possible roles in cassava development and especially the development of PPD.

1.7 Genetic transformation of cassava

Genetic transformation refers to transferring gene(s) to cells competent for regeneration, subsequent selection of transformed tissue, and regeneration of plants from cells containing the transferred genes in the genome. Genetic transformation can increase the genetic diversity available to crop breeders by permitting the use of genes from other species, or by altering the expression level of native genes. It is a powerful tool for understanding complex biochemical pathways, since it allows researchers to knock out certain genes using antisense technique, sense gene silencing or over-express specific genes and compare the phenotypes of the transgenic and non-transformed plants. It has also made it possible for researchers to analyse gene regulation by expressing reporter genes driven by a promoter or/and its deletion series.

To establish a plant genetic transformation system two main factors are required (Potrykus, 1991). Firstly, an efficient *in vitro* culture system needs to be developed to

regenerate plants from transformed tissues. Secondly, a technique for efficient transfer and stable integration of transgenes into the plant genome is essential and methods for identifying and selecting for transformed cells are also required.

The availability of cassava transformation techniques provides opportunities for cassava researchers to understanding the unique metabolism and molecular biology of this plant and seeking for the possibilities to improve quality and quantity of cassava. Concerted efforts to develop a genetic transformation system for cassava were initiated in the early 1990s but the initial progress in cassava transformation was slow due to limited funding and subsequent limited research input, technical difficulties of the crop *in vitro*, and the difficulties in selecting and regenerating plants from transformed cells. In 1996, breakthroughs were made in cassava transformation and transgenic cassava plants were produced using either *Agrobacterium* or particle bombardment in manipulating the embryogenic culture systems (Li *et al.*, 1996; Schopke *et al.*, 1996).

1.7.1 Tissue culture and regeneration of cassava

Somatic embryogenesis in cassava was first reported in 1982 (Stamp and Henshaw) and is now the most commonly used regeneration method for cassava. This system involved the production of primary somatic embryos, usually from leaf or meristem explants or zygotic embryos on MS medium supplemented with sucrose and auxins such as 2,4-D or picloram. The competence for somatic embryogenesis in cassava is restricted to meristematic and embryonic tissues, and somatic embryos can only be induced on a limited number of explants, e.g. cotyledons or embryonic axes from zygotic embryos (Stamp and Henshaw, 1982, Konan *et al.,* 1994), immature leaf lobes (Stamp and Henshaw, 1987; Szbados *et al.*, 1987; Matthews *et al.*, 1993; Raemakers, 1993;), meristems and shoot tips (Szabados *et al.*, 1987; Narayanaswami *et al.*, 1995; Puonti-Kaerlas *et al.*, 1997) and immature inflorescences (Mukherjee, 1995) on media containing auxin.

The embryogenic potential in cassava is highly genotype dependent. Some cultivars such as TMS 60444 generate somatic embryos with ease, while with some cultivars no somatic embryos could be obtained despite repeated efforts. Some techniques, such as

treatment with 2,4-D (Matsumoto *et al.*, 1991; Raemakers, 1993) and addition of ABA (Konan *et al.*, 1994) or supplementary cupric sulphate to the embryo induction medium (Schopke *et al.*, 1993), were shown to be effective in improving the embryo induction. Induction of somatic embryogenesis has been extended to several cultivars due to the improvement of the original protocol (Szabados *et al.*, 1987; Raemakers *et al.*, 1993; Taylor *et al.*, 1997).

When somatic embryos were maintained on auxin-containing medium, some of them spontaneously started developing to torpedo-shaped and eventually mature embryos with greening cotyledons, which could be enhanced if the embryos were transferred to a medium containing low amounts of BA and auxin or hormone-free medium (Stamp and Henshaw, 1982, 1987). The efficiency of secondary somatic embryogenesis had been shown to be higher in liquid cultures than on solid cultures (Raemakers *et al.*, 1993; Li *et al.*, 1995). The germination frequency of mature somatic embryos was usually low and the root development was incomplete. Use of NAA allowed efficient production of somatic embryos with high germination capacity (Taylor *et al.*, 1996).

It was observed that a new tissue type consisting mainly of small globular embryo-like structures was generated from a fraction of the cycling somatic embryos maintained on MS medium supplemented with 12 mg/l picloram. When isolated, these produced a highly friable embryogenic callus (FEC) (Taylor *et al.*, 1996). By substituting the MS medium with GD medium (Gresshoff and Doy, 1974) the frequency of FEC production could be increased. Pure FEC could be used to establish a rapidly proliferating embryogenic suspension in liquid SH (Schenk and Hildebrandt, 1972) medium supplemented with 6% sucrose and 12 mg/l picloram. Cotyledon explants derived from cycling somatic embryos showed the highest competence for organogenesis, while those from primary somatic embryos responded very poorly.

1.7.2 Cassava transformation systems

Research on genetic transformation of cassava had been focused on two methods: biolistic and *Agrobacterium tumefaciens* mediated gene transfer.

The biolistic method has been used successfully to transform both agricultural and non-agricultural plants, animal cells, insect and fish embryos, algae fungi, pollen, bacteria, and intracellular organelles. It was found to be an effective method for transfer genes into monocots such as rice, wheat and maize. Transgenic plants have been obtained in a number of plant species by the biolistic method, such as rice, barley and wheat.

The cassava transformation system using particle bombardment was based on the establishment of an embryogenic suspension culture system. Transformation via particle bombardment was initially tried on a morphogenic culture system of somatic embryogenesis, which resulted only in chimeric embryos (Schopke *et al.*, 1993). Efforts were then focused on the development of friable embryogenic callus and embryogenic suspension cultures, which were proven to be suitable tissues for transformation using particle bombardment in other plant systems (Vasil, 1995; Christou, 1992). Gresshoff and Doy basal medium supplemented with 4-amino-3, 5, 6, trichloro-picolinic acid (picloram) induced successfully the formation of friable embryogenic callus (Taylor *et al.*, 1996). From this callus, highly totipotent embryogenic suspension cultures were established and transgenic cassava containing *npt*II and *uid*A genes were generated by bombarding these cultures with gold particles coated with a plasmid containing these two genes (Schopke *et al.*, 1996). At the same time, a group in Wageningen used the same approach with luciferase marker gene and obtained transgenic plants (Raemakers *et al.*, 1996).

Gene transfer using *Agrobacterium tumefaciens* is well-established in a number of dicot plants. This transformation system takes advantage of the nature of *Agrobacterium tumefaciens* causing crown gall disease on plants, which can infect wounded plant tissue, transfer part of its T-DNA into plant cells and integrate the T-DNA into plant genome. During the establishment of *Agrobacterium*-based cassava transformation system, four different *Agrobacterium* strains were initially tested to identify the ones leading to the highest transient expression rates (Li *et al.*, 1996). Two strains LBA4404 (pTOK233) and LBA4044 (pBin9GusInt) showed the highest transient transformation rates and were selected for stable transformation. Cotyledon explants were harvested from secondary somatic embryos, inoculated with the *Agrobacterium* and then transferred onto selection medium. Shoot regeneration was induced by cytokinin

benzylaminopurine. Transgenic cassava plants were successfully regenerated from antibiotic resistant callus (Li *et al.*, 1996).

Most recently, a novel approach based on biolistics and regeneration via organogenesis was developed for genetic transformation of cassava (Zhang *et al.,* 2000). Green cotyledons from germinating somatic embryos were used as target tissue and CaMV 35S promoter-GUS was transferred to the tissue by particle bombardment. By optimising parameters, a set of bombardment conditions was evaluated and the most efficient condition was identified. The optimal condition was 0.5 μg plasmid DNA per shot, 12.5 cm flying distances between the filter and the target, and 20 hours for both pre- and post- plasmolysis time. The use of a low concentration of BA (0.4 mg/l) in the shoot elongation medium following initial shoot induciton on a medium containing 1 mg/l BA, appeared to improve the regeneration and subsequently increased transformation frequency.

With the establishment of transformation systems, focus is shifted to use the systems for analysis of promoters and genetic manipulation of cassava. However, cassava transformation is still a labour- and time-consuming process, on which further optimisation could be done. Another drawback of the current transformation systems is their cultivar-dependency, which needs to be addressed should the techniques to be employed to improve the productivity of cassava in the tropical countries.

1.7.3 Application of cassava transformation

Since the first reported transgenic cassava plants (Schopke *et al.,* 1996; Li *et al.,* 1996), much effort was spent on improving and optimizing transformation systems (Munyikwa *et al.,* 1998), and transformation with genes of agronomic interests.

One of the focuses is to improve the resistance of the crop to diseases and pests. Gene constructs containing the coat protein gene of cassava common mosaic virus, the AC1 (replication-associated protein) gene and defective interfering (DI) genes of African cassava mosaic virus, viral antisense RNAs and truncated viral gene products were transferred into cassava respectively and transgenic plants containing these genes were

produced (Schopke *et al.*, 1998; Taylor *et al.*, 1998; Scharer-Hernandez *et al.*, 1998) and the effects of these transgenes on the resistance of cassava to mosaic virus diseases are being assessed. Bacterial blight resistance gene *Xa*21 from rice, which had been shown to be able to confer resistance to *Xanthomonas oryzae* pv *oryzae* (bacterial blight) when transferred into rice (Zhang *et al.*, 1998), was also introduced into cassava to study its effect on the resistance to cassava bacterial blight (Taylor *et al.*, 1998). Two groups of researchers had transferred *cry*1A(b) gene of *Bacillus thuringiensis* into cassava in the hope that transgenic cassava resistant to stem borer could be obtained (Legris *et al.*, 1998; Chavarriaga *et al.*, 1998).

Attention was also drawn to the quality improvement of cassava. Genetic manipulation of cyanide toxicity was attempted by transferring a hydroxynitrile lyase cDNA, which catalyses the breakdown of cyanohydrin, a product of deglycosylation of linamarin (Arias-Garzon and Sayre, 1998). Most recently, three lines of transgenic cassava plants resistant to herbicide Basta were produced by *Agrobacterium*-mediated transformation (Sarria *et al.*, 2000).

1.7.4 The search for promoters for genetic modification of cassava

A few promoters have been used in cassava transformation. CaMV 35S was successfully used in the initial optimization of the cassava transformation systems (Li *et al.*, 1996; Schopke *et al.*, 1996) and genetic manipulation of cassava cyanide toxicity and resistance to pests (Arias-Garzon and Sayre, 1998; Legris *et al.*, 1998; Chavarriaga *et al.*, 1998). A cassava virus promoter from vein mosaic virus (CsVMV) was isolated and this promoter (including its derivatives) showed constitutive, root- and phloem-specific activities in tobacco and rice (Verdaguer *et al.*, 1998). These promoters were fused to the reporter gene *uidA* and transferred into cassava to test their activities (Taylor *et al.*, 1999). Li *et al.* (1998) used a senescence specific promoter *sag* from *Arabidopsis* to drive a *ipt* gene controlling cytokinin biosynthesis from *A. tumefaciens* in order to prolong cassava leaf life (Li *et al.*, 1998).

Although constitutive promoters or tissue- and developmental- specific promoters as mentioned above can be used in genetic manipulation of cassava, these promoters may

not work efficiently in certain aspects of manipulation. For example, PPD is a special problem for cassava crop production, a PPD-specific promoter would be required for efficient manipulation of PPD.

A β-glycosidase gene promoter from cassava was shown to be root-specific by analysing the gene expression and the transient activity of the promoter in different tissues (Liddle *et al.,* 1997). The promoter of a gene encoding protein synthesis elongation factor in cassava was isolated from a genomic library and three putative control elements including TEF1, TELO and TATA boxes were identified (Suhandono *et al.,* 1998). Using particle bombardment, it was shown that the promoter was transiently active in leaves of cassava, *Arabidopsis*, radish and pea (Suhandono *et al.,* 1998). Searching for root-specific promoters in cassava has also been carried out. Various promoters were tested for tissue specific activity in cassava by transient assays and a low frequency of transient gene expression was observed in the storage roots (Arias-Garzon and Sayre, 1993). A number of cDNA clones showing root specific expression were isolated by a differential screen and the isolation and characterization of their corresponding genomic DNA sequences would help to identify root-specific promoters (Huang *et al.,* 1998; Bohl-Zenger *et al.,* 1998).

1.8 The aims of the project and strategy

It is known that the activity of PAL increased dramatically during PPD, which appeared to be related to rapid accumulation of phenolic compounds such as scopoletin that is suggested to be related to the development of PPD (Wheatley and Schwabe, 1985; Rickard, 1982). These, together with the roles of PAL genes in phenylpropanoid metabolism and plant defense responses in other well-studied plant systems, suggest the possible roles of PAL in cassava. In order to understand the role of PAL in cassava during normal development and in response to stresses, particularly PPD, this project was proposed to investigate the spatial and temporal expression patterns of a PAL gene. A partial PAL gene was cloned by Pereira *et al* (1999) and two cDNA clones were isolated by Han (2000). However, limited expression data were obtained, especially spatial expression patterns, mainly due to the nature of cassava tissues. It was, therefore, proposed to study the PAL gene expression using promoter-GUS fusion in

transgenic plants, taking advantage of the newly developed cassava transformation system. To this end a genomic PAL clone was to be isolated. The promoter would then be analysed at sequence level for putative regulatory domains or elements by comparing its sequence with well-characterised PAL promoters in other plants. Based on the properties of the promoter sequence, promoter deletions and mutations would be made and fused to GUS reporter gene to form different constructs. These constructs would be tested for the transient activities of the promoter in cassava suspension cells and *in vitro* leaves before they would be transformed into cassava to obtain transgenic plants for stable GUS analysis. There are a number of examples on the activity of dicot PAL promoters in heterologous dicot hosts and one case of a monocot (rice) PAL promoter in a dicot (tobacco). Yet there is no evidence that a dicot PAL promoter is active in a monocot. Therefore, it was proposed that cassava promoter constructs be tested in rice. Since rice transformation is much more efficient and less time consuming, it was thought that GUS expression data could be obtained from transgenic rice well before transgenic cassava plants were obtained if both series of transformation were started together, which would give a preview of the expression pattern of PAL promoter.

Ethylene is involved in signalling during stresses such as wounding in other plant systems as described in section 1.6 and it has been suggested that wounding and water loss were associated with the development of PPD in cassava. It was shown that there was an increase in ethylene production during PPD but exogenous application to storage roots did not affect PPD (Plumbley *et al.*, 1981; Hirose *et al.*, 1984). Therefore, the role of ethylene in PPD is not clear. One way to understanding the possible role of ethylene in PPD is to study the expression of genes related to the biosynthesis of ethylene, such as ACC oxidase (ACO) genes. An ACO cDNA had been isolated from a PPD-related cDNA library of cassava (Han, 2000). This indicated the expression of an ACO gene during PPD, which was in line with the ethylene production during PPD. However, it was not clear how this gene was regulated during PPD as no expression data was obtained due to the relatively low abundance of its transcripts. Therefore, it was proposed to study ACO genes in a similar approach as for PAL gene, isolating and characterizing ACO promoter. The aim is to understand how the ACO gene is regulated during cassava development, especially during the development of PPD, and to

investigate the possibility of genetic manipulation of PPD with antisense- or sense-gene silencing using the ACO gene.

CHAPTER TWO: MATERIALS AND METHODS

2.1 Plant materials

Cassava cultivars M NGA 1 and TMS60444, and transgenic cassava plants were grown in the greenhouse, at temperatures of 27°C during the day time and 15°C during the night time, with a photoperiod of at least 16 hours a day. Levington (Fisons) M2 compost was used for growing cassava. *In vitro* cassava was grown in a growth chamber with a 16 hours photoperiod at 28°C in the light and 15°C in the dark. Preparation of plant materials for transient and stable GUS assays of MePAL2 promoters and transformations were as described in Sections 2.7 and 2.8

Rice plants (cultivar *Oryza sativa* L. Japonica cv. Taipei 309 (TP309) and transgenic rice were grown in 13 cm pots in a greenhouse, with a 14 hour light period at 30°C and in the dark at 21°C per day. *In vitro* rice plantlets were grown in a growth chamber with a 16 hours photoperiod at 28°C in the light and 15°C in a dark a day.

2.2 DNA extraction

2.2.1 Plasmid DNA extraction

Three different methods of plasmid extraction were used according to the downstream applications. For checking recombinant plasmids, the mini-preparation method was used. Plasmid DNA used for automatic sequencing was prepared with the Wizard Plus SV Minipreps DNA Purification System (Promega). Plasmid DNA used for plant transformation was purified on a large scale and at high concentration using the Qiagen Midi Plasmid Purification system (QIAGEN).

Mini-preparation of plasmid DNA:

A single bacterial colony was inoculated into 2 ml of LB liquid medium (10 g tryptone, 10 g NaCl, 5 g yeast extract in one litre water, pH to 7.0 with NaOH and autoclaved) containing the appropriate antibiotic (such as ampicillin) in a loosely capped 15 ml tube. The bacteria were grown overnight at 37°C and 150 rpm in a shaker. The overnight culture (1.5 ml) was transferred into a microcentrifuge tube and centrifuged at 13000 rpm for 30 seconds at 4°C. After the supernatant was removed the pellet was resuspended in 100 µl of ice-cold Solution I (50 mM glucose, 25 mM Tris·Cl (pH 8.0), 10 mM EDTA (pH 8.0), (autoclaved and stored at 4°C) by vigorously vortexing. Freshly prepared Solution II (200 µl) (0.2 N NaOH and 1% SDS) was then added and the content was mixed by inverting the tube gently five times. Ice-cold Solution III (150 µl) (for 100 ml of the solution: 60 ml of 5 M potassium acetate; 11.5 ml glacial acetic acid and 28.5 ml of H_2O) was added to the tube. The tube was gently inverted a few times, placed in an inverted position for 10 seconds and kept on ice for 3-5 minutes. The mixture was centrifuged at 13000 rpm for 5 minutes at 4°C and then the supernatant was transferred to a fresh tube. The supernatant was extracted with phenol:chloroform: isoamyl (25:24:1) once and precipitated with 2 volumes of 100% ethanol. Precipitated DNA was washed with 1 ml ice-cold 70% ethanol. The DNA pellet was air-dried and resuspended in 50 µl MilliQ water containing DNAase-free RNase (20 µg/ml) (Sigma, UK). Then the plasmid DNA was checked by restriction digestion.

Wizard plus SV Minipreps DNA Purification System (Promega):

Five ml of an overnight plasmid-containing *E.coli* culture was centrifuged for 5 minutes at 4000 rpm (Sorvall SS34). The supernatant was discarded and 250 µl of Cell Resuspension Solution was added to the pellet. The pellet was dissolved by vortexing and transferred to a 1.5 ml microcentrifuge tube. Cell Lysis Solution (250 µl) was then added, mixed by inverting the tube 4 times, and incubated for 5 minutes until the cell suspension cleared. Alkaline Protease Solution (10 µl) was added and incubated for 5 minutes at room temperature. Neutralization Solution (350 µl) was added to the tube and the mixture was then immediately mixed by inverting the tube 4 times until bacterial lysate appeared. Then the bacterial lysate was centrifuged at 13,000 rpm in a microcentrifuge for 10 minutes at room temperature. The cleared lysate was transferred

to a Minipreps Spin Column and centrifuged at 13000 rpm for 1 minute at room temperature. The flow-through was discarded from the Collection Tube. The Minipreps Column was washed twice with 750 µl and 250 µl Wash solution respectively. The Minipreps Spin Column was then transferred to a 1.5 ml microcentrifuge tube and plasmid DNA was eluted with 100 µl of Nuclease-Free Water.

Midi Pasmid Purification (QIAGEN):
An overnight *E.coli* culture (50 to 100 ml) was centrifuged for 10 minutes at 4000 rpm and the pellet was resuspended in 4 ml of buffer P1. Four ml of buffer P2 was added to the tube. The contents were mixed gently and incubated at room temperature for 5 minutes. Four ml of chilled buffer P3 was added and mixed gently. The mixture was incubated on ice for 15 minutes and centrifuged at 20,000 g for 30 minutes at 4°C. The supernatant was transferred to a QIAGEN-tip 100, which had been equilibrated with 4 ml of QB buffer. After the tip was drained, it was washed twice with 10 ml of buffer QC. The DNA in the tip then was eluted with 5 ml of buffer QF. The eluted DNA was precipitated with 3.5 ml isopropanol and centrifuged immediately. The supernatant was carefully removed and DNA pellet was washed twice with 2 ml of 70% ethanol. After the pellet was air-dried, the DNA was resuspended with 100 µl of MilliQ water.

2.2.2 λ phage DNA extraction

Lambda plaques were cored out and each was put into a microcentrifuge tube containing 0.5 ml SM buffer (for 1 litre of SM buffer 5.8 g of NaCl, 2 g of $MgSO_4 \cdot 7H_2O$, 6.05 g of Tris base and 5 ml of Gelatin were added and then adjusted to pH 7.5 with HCl before autoclaving). To amplify phage particles, 50 µl of the lambda in SM buffer was grown on LB-agar plate as described in Section 2.4.1. After the phage plate was kept in a cold room for 1 or 2 hours, 5 ml of SM buffer was added to the plate, which was then shaken slightly in the cold room for two hours to obtain a plate lysate. Then the plate lysate was transferred into a few microcentrifuge tubes and used in bulk liquid lysate later on or stored in 4°C by adding a drop of chloroform.

One ml of overnight culture *E.coli* LE392 (host for the phage) grown in NZYCM (for 1 litre of medium, 5 g of NaCl, 5 g of bacto-yeast extract, 2 g $MgSO_4$, 11 g Casein were added and then adjusted to pH 7 with NaOH before autoclaving) was pipetted into 150 ml NZYCM in 250 ml flasks and incubated at 37°C, 120 rpm until OD_{600} = 0.5-0.6, then 1 ml of the plate lysate was added and the incubation was continued. After 4-5 hours, when lysis happened, the lysate was centrifuged at 10,000 rpm (Sorvall GSA) for 15 minutes at 4°C. The supernatant was then transferred to a flask. Fifty µl of RNase (10 mg/ml) and 50 µl DNase (10 mg/ml) were added to a final concentration of 1 µg/ml to the flask with supernatant to digest bacterial DNA and RNA. The digestion was carried out at 37°C for 1 hour. To each sample 9 g of NaCl (final concentration 1 M) was added and dissolved into the supernatant and the flasks were kept on ice for 1 hour. To each sample 15 g (final concentration 10% w/v) PEG 8,000 was added and dissolved slowly using a stirrer. The flasks were kept on ice overnight to precipitate phage particles. After centrifugation at 10,000 rpm (GSA Sorvall) for 15 minutes at 4°C, the supernatant was discarded and any drops of fluid in the tube were removed. Then 4 ml of SM buffer was added to the pellet, which was resuspended slowly and thoroughly. Four ml of chloroform was added to it and vortexed gently to form an emulsion. After centrifugation at 3000 rpm (GSA Sorvall) for 15 minutes at 4°C, the aqueous phase was transferred to another tube. 160 µl of EDTA (1.5 M pH = 8.5) (to a final concentration of 20 mM), 5 µg proteinase K (final concentration 50 µg/ml) and 200 µl of 10% SDS (final concentration 0.5%) were added to each sample. After being mixed by inversion and incubated for 1 hour at 56°C, the mixtures were extracted once with 1 volume phenol (equilibrated with 0.5 M Tris-HCl pH 8), twice with 1 volume phenol:chloroform: isoamyl (25:24:1) and once with chloroform: isoamyl alcohol. Then 0.1 volume of sodium acetate (3 M pH7.0) and 0.7 volume of isopropanol were added and kept at -20°C for 30 minutes to precipitate the λ DNA. After centrifugation at 14,000 rpm (GSA Sorvall) for 30 minutes at 4°C, the pellet was washed with 10 ml 70% ethanol three times. The air-dried λ DNA pellet was resuspended in 500 µl TE buffer (1.21 g of Tris base and 0.37 g of EDTA in 1 litre, adjusted to pH 8.0 with HCl before autoclaving) and stored at -20°C.

2.2.3 Genomic DNA extraction

Young leaf (cassava or rice) tissue (7-8 g) was ground in liquid nitrogen to a fine powder. The powder was distributed into two tubes (4 g powder for each tube), to which 15 ml pre-warmed (40-50°C) Dellaporta extraction buffer (100 mM of Tris-HCl (pH 8.0), 50 mM EDTA, 500 mM NaCl, to which 0.07% v/v β-mercaptoethanol and 1% w/v of PVP were added to the buffer just before use) was added. Then 1 ml 20% SDS (w/v) was added to each tube. The mixture was incubated in a shaking water bath at 65°C for 10 minutes, during which the tubes were inverted every 2 minutes. Then 5 µl of 5 M potassium acetate was added and mixed vigorously. The mixture was centrifuged at 15,000 rpm (Sorvall SS34) for 20 minutes at 4°C after which it was incubated on ice for 40-50 minutes. The supernatant was filtered through miracloth (or a double thickness of gauze) into a tube containing 10 ml of isopropanol precooled at -20°C. The tubes were gently inverted a few times to mix and then kept at -20°C overnight. The DNA was precipitated by centrifugation at 15,000 rpm (Sorvall SS34) for 15 minutes at 4°C and the pellet was washed three times with 2 ml 70% ethanol. After the pellet was air-dried at room temperature, it was redissolved in 700 µl TE buffer (sometimes incubation at 65°C for 5-10 minutes was carried out to improve the process). Then the mixture was transferred to a new 1.5 ml microcentrifuge tube and centrifuged at 12,000 rpm (Sorvall SS34) for 10 minutes at room temperature in a microcentrifuge. The supernatant was transferred to a new 1.5 ml microcentrifuge tube and DNase-free RNase was added at a final concentration of 10 µg/ml, and the tube was incubated in a water bath at 37°C for 30 minutes. The supernatant was extracted with 1 volume of phenol:chloroform: isoamyl (25:24:1). To the aqueous phase 75 µl 3 M sodium acetate and 500 µl isopropanol (pre-cooled) were added to the tube and then kept at -20°C overnight. Then the tube was centrifuged at 15,000 rpm (Sorvall SS34) for 15 minutes at 4°C. The supernatant was discarded and the pellet was washed three times with 500 µl 70% ethanol (pre-cooled at -20°C) and vacuum dried. The DNA was dissolved in 200 µl sterile water and stored at -20°C. The concentration of the genomic DNA was measured by electrophoreses in an agarose gel and by spectrophotometry.

2.3 DNA manipulation

2.3.1 Genomic, plasmids, phage or DNA fragment digestion

Plasmids, phage DNA or DNA fragments were mixed with a certain volume of appropriate 10 x buffer according to the instruction of manufacturers and MilliQ water to create 1 x buffer condition suitable for the restriction enzyme. The restriction enzyme was added to the mixture according to the quantity of DNA and activity of the enzyme. The digestion mixture was then incubated at the appropriate the temperature (usually at 37°C) for the enzyme as suggested by the manufacturers.

For genomic DNA, 10 μg of DNA for each digestion was used unless otherwise stated. To the DNA, 5 μl of 10 x buffer and 1.1 μl of 0.1 M spermidine were added, then MilliQ water was added to final volume 48 μl and finally the restriction enzyme 2 μl (12 units/μl) was added. The digestions were incubated overnight at the appropriate the temperature (usually at 37°C) for the enzyme as suggested by the manufacturers.

2.3.2 PCR amplification

Synthesised primers (see primer designing) based on the DNA template were used for PCR amplification. For each 20 μl PCR reaction 10 ng template DNA was used. The PCR reactions were started with one cycle of 3 minutes denaturation at 94°C. Then followed by 35 cycles of denaturation for 1 minutes at 94°C, primer annealing for 1 minute at 53°C (depending on the annealing temperature of the primers), and DNA synthesis for 3 minutes at 72°C, finally completed with 72°C for 10 minutes. The reaction products were electrophoresed in an agarose gel and visualized under UV light.

2.3.3 DNA fragment / band purification

PCR products or fragments from restriction enzyme digestion were separated in 0.5 x TBE agarose gel (108 g of Tris base, 55 g of Boric acid and 9.3 g EDTA-disodium salt in 1 litre for 10 x TBE buffer) and purified from the gel using a SephaglasTM Bandprep kit from Pharmacia Biotech. The DNA band to be purified was cut out from the gel and placed in a microcentrifuge tube. Gel Solubilizer 1 μl/mg gel was added to the tube. To enhance the DNA extraction efficiency from the gel, 5 μl 50% glacial acetic acid was added to every 250 μl Gel Solubilizer. The mixture was incubated at 60°C for 5-10 minutes until the gel was dissolved completely. Sephaglas BP, 5 μl for each μg of DNA, was added and vortexed gently. The mixture was incubated at room temperature for 5 minutes and pulse-centrifuged for 30 seconds. The supernatant was discarded and the pellet was washed three times with 80 μl of 70% ethanol for every 5 μl of Sephaglas BP and then air-dried. TE buffer was added to the pellet and resuspended to elute the DNA. The mixture was incubated at room temperature for 5 minutes with periodic agitation. After 30 seconds centrifugation, the supernatant was transferred to fresh tube, stored at -20°C.

2.3.4 Recombinant DNA techniques

2.3.4.1 Sub-cloning DNA fragment into vectors

Lambda phage DNA or DNA fragments were digested with appropriate restriction enzymes to release appropriate bands for subcloning. pUC19 was used as a vector in most experiments. Ligation of the inserts into pUC19 was carried out at a molar ratio of 3 to 1 between insert DNA and vector DNA. The ligation enzyme was T4 DNA ligase. The total volume was 10 μl and DNA (vector and insert together) was less then 100 ng. Ligation was buffered with 1 x ligation buffer and incubated at 12°C overnight. Then ligation mixture was transformed into fresh competent cells of *E.coli* DH5α prepared using calcium chloride.

2.3.4.2 Preparation of competent *E.coli* cells and transformation

E.coli DH5α (ϕ80*dlacZ*ΔM15, *recA*1, *endA*1, *gyr A*95, *thi*-1, *hsd R*17 (r_K^-, m_K^+), *supE*44, *relA*1, *deoR*,Δ(*lac ZYA-argf*) U169) was used as the host cell for plasmid transformation.

A single colony of DH5α was picked from an overnight plate culture at 37°C and inoculated into 10 ml of LB medium (10 g of Tryptone; 5 g of Yeast extract and 10 g of NaCl in 1 litre water, adjusted to pH to 7.0 with NaOH and autoclaved). The culture was incubated at 37°C with vigorous shaking (150 rpm) overnight. One ml of the overnight culture was transferred to 50 ml pre-warmed LB medium in a 250 ml conical flask and the culture was incubated at 37°C with shaking at 150 rpm. When the OD_{600} of the culture reached 0.5-0.6, the culture was aseptically transferred to a sterilized and ice-cold centrifuge tube and kept on ice for 10 minutes. The cells were recovered by centrifuging at 4000 rpm (Sorvall SS34) for 10 minutes at 4°C. The supernatant was decanted and the tube was stood in an inverted position for 1 minute to allow the last traces of media to drain. The pellet was resuspended in 10 ml of ice-cold 0.1 M $CaCl_2$ and stored on ice. The cell suspension was centrifuged again as above. The supernatant was discarded and the pellet was resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$. 200 µl of competent cells was transferred to a sterile microcentrifuge tube using a chilled sterile pipette tip, and 10 µl recombinant DNA (no more than 50 ng) from ligation reaction was added. Uncut pUC19 vector DNA was used as a positive control, whereas 200 µl competent cells without any DNA or restriction digested pUC19 as a negative control. The cells and DNA were mixed by swirling gently and stored on ice for at last 30 minutes. The mixture was then heat-shocked by placing the tubes in a 42°C water-bath for precisely 90 seconds. Immediately after heat-shock the tubes were replaced in the ice for 10 minutes. Then 800 µl LB medium was added to the cells and incubated at 37°C in a shaking water-bath for one hour to allow expression of the antibiotic resistance gene. The cells (150 µl) were spread onto a LB-agar plate containing 50 µg/ml ampicillin, 20 µg/ml X-gal, 20 µg/ml IPTG. The plates were incubated at 37°C overnight.

White colonies on each plate were chosen to inoculate another LB plate containing ampicillin, IPTG and X-gal and the plate was incubated overnight to confirm that the white colonies were genuine recombinants. Three positives from each transformation events were chosen for plasmid extraction and restriction digestion for further confirmation.

2.3.4.3 Confirmation of the amplified recombinant DNA

Recombinant DNA was tested by either PCR amplification, or restriction digestion, or sequencing.

PCR amplification: Three positives from each transformation were grown in a LB plate with appropriate antibiotics overnight at 37°C. The colonies were picked and transferred a tube containing 100 µl of MilliQ. The bacteria were boiled for 1 minute, centrifuged for 2 minutes at 13,000 rpm and the supernatant was used as template for PCR. PCR was performed using appropriate primers for the insert of DNA and 2 µl of supernatant extracted from boiled *E.coli* as template. If the band of PCR product was exactly the same size as positive control after running in a TBE gel, the clone was considered as positive.

Plasmid extraction and restriction digestion: Plasmid DNA was extracted from 4-6 positive clones and digestions performed using the appropriate restriction enzyme. The digested mixture was separated in a TBE agarose gel to check whether expected fragment was released by the digestion.

Sequence confirmation: PAL promoter constructs were sequenced to make sure they had the correct sequences before being used for transient assays of the PAL promoter and for plant transformation.

2.3.5 Southern hybridization

2.3.5.1 Southern blotting

Restriction-digested plasmid DNA, or lambda DNA, or genomic DNA, and size marker *Hin*dIII-cut λ DNA, were run in 0.8% agarose gel, at 20 volts (1 v per 1 cm) overnight without adding ethidium bromide (EtBr). Then the gel was stained for 30 minutes in EtBr solution (100 μg /1 liter). Afterwards the gel was soaked in 0.25 M HCl for 15 minutes, denatured in 1.5 M NaCl + 1.5 N NaOH with gentle shaking for 30 minutes and neutralized in 1 M Tris-HCl (pH 7.4) + 1.5 M NaCl for 30 minutes. The gel was then rinsed with distilled water and trimmed. Finally the gel was blotted onto Hybond N^+ overnight as described (Sambrook *et al.*, 1989). The membrane was rinsed in 2 x SSC (for 1 litre of 20 x SSC: 175.32 g of NaCl and 88.2 of sodium citrate, adjusted to pH 7.0) for 5 minutes, air-dried briefly and vacuum-fixed at 80°C for 2 hours.

2.3.5.2 Probe preparation

Purified DNA fragments were used to prepare probes using an Oligolabelling kit from Pharmacia Biotech. Template DNA (30-50 ng) was denatured at 100°C for 5 minutes and kept in ice for 3 minutes. Then 10 μl Reagent mix, 2-5 μl (α $^{-32}$P) dCTP was added, and ice-cold MilliQ water was added to a final total volume of 49 μl. Klenow Fragment 1 μl (5 units/μl) was added to the mixture. The labelling reaction mixture was incubated at 37°C for 30 minutes.

2.3.5.3 Hybridization Procedures

Filters were placed in 50 -100 ml preheated pre-hybridization buffer (6 x SSC; 1% SDS; 5 x Denhardt's solution (for 50x of 500 ml Denhardts solution, 5 g of ficoll, 5 g of polyvinylpyrrolidone and 5 g of bovine serum albumin were added to MilliQ water and stored to -20°) in a hybridization sandwich box. Sufficient sonicated heterologous DNA (herring sperm DNA) was denatured first in a boiling water-bath for three minutes and then added to the pre-hybridization solution to give a final concentration of

100 μg/ml. The filter was prehybridized by incubating at 60°C with slight shaking for 1-2 hours in a sandwich box in a shaking incubator. The radio-labelled probe was added after being denatured at 95°C for 5 minutes. Hybridizitions were carried out overnight at the relevant temperature. Following hybridization, the filters were washed twice with 2 x SSC, 0.1% (w/v) SDS at room temperature for 10 minutes. Then washed twice in 1 x SSC, 0.1% SDS at a particular temperature for 10 to 15 minutes. The filters were sometimes washed with higher stringency solutions such as 0.1 x SSC, 0.1% SDS. The filters were then wrapped in cling film, and X-ray films were exposed to the filters between two intensifying screens at -80°C overnight or for several days depending on the strength of the signals.

2.4 Screening cassava genomic DNA library and subcloning

2.4.1Titering genomic library

A cassava genomic library constructed from cultivar M BRA 534 (a gift from Professor M.A. Hughes, University of Newcastle), was used to screen for PAL and ACC oxidase genes. This library consisted of *Sau*3A partially-digested cassava genomic DNA cloned in the *Bam*HI site of λ EMBL3,

E.coli strain LE392 (*hsdR*574, (r_K^-, m_K^+), *sup E*44, *sup F*58, *lac Y*1 or Δ(*lacIZY*)6, *galK*2, *galT*22, *met B*1, *trp R*55) was used as the host for lambda phage. A single colony of LE392 was inoculated into 10 ml LB medium at 37°C 120 rpm overnight. Next day 1 ml of the overnight culture was pipetted into 50 ml LB medium containing 0.2% maltose and incubated at 37°C 120 rpm until OD_{600}=0.5-0.6. After centrifuging for10 minutes at 4°C, 15 ml ice-cold 10 mM $MgSO_4$ was added to the pellet, vortexed and the plating cells were kept on ice for use.

A series of dilutions of the library were made to titre the library. Five μl of genomic library phage mixture was added to 45 μl of SM buffer (10^{-1} dilution), and mixed well. Then 30 μl of the 10^{-1} dilution was added into 270 μl SM buffer and mixed well (10^{-2} dilution). Similarly a series of dilutions down to 10^{-8} dilution was prepared. Diluted phage 100 μl (from 10^{-4} -- 10^{-8}) was mixed with 100 μl plating cells LE392

(OD_{600}=0.5) in a 5 ml Bijoux and incubated for 30 minutes at 37°C. Then 3 ml molten (45°C) LB top agar (see media) was poured into the Bijou and immediately poured onto LB plates pre-warmed at 37 - 40°C. After the top agar had set and slightly air-dried, the plates were incubated in an inverted position at 37°C overnight.

2.4.2 Screening cassava genomic library

The number of phages from the library required to get target gene clones was estimated according to the following formula (Brown, 1990).

$N = \ln (1-P)/\ln (1-\delta)$

N : number of plaques to be screened

P : probability to get the target clone(s)

δ : proportion of genome in insert

The probability was set at 99% (P=99%). The sizes of the insert and cassava genome were 15 kb and 772 Mbp (Awoleye, 1994), so $\delta = 15/772 \times 10^3$. Then the number of plaques required for obtaining the target clone at 99% probability was $N = 2.37 \times 10^6$ plaques. As the titre of the library is 2.1×10^8 pfu/ml, therefore 10 µl phage mixture from the library was used for the screen.

To 10 µl phage of the library, 990 µl SM buffer was added and mixed well. Then 200 µl of the diluted phage mixture was added to 200 µl LE392 plating cells. After incubation at 37°C for 30 minutes, molten top agarose (0.8% agarose in LB) (45°C – 47°C) was poured on to the mixture and immediately poured onto a LB-agar plate. After overnight incubation at 37°C, the plates were placed at 4°C for 1.5 hours prior to performing the plaque lifts.

The Hybond N^+ (Amersham) filter was gently placed on one edge of the phage plate and progressively laid down on the surface of the agar. The first filter was left for 1 minute at room temperature and a duplicate filter for 5 minutes. To record the orientation of the filter on the plate, a 20-gauge needle was used to puncture through the filter into the agar medium asymmetrically at 3-5 points around the edge of the

plate. Using a blunt forceps, the filter was lifted slowly from the plate and placed plaque side up on a Whatman paper. The filter was then processed through three trays. Three trays were set up, each containing a sheet of Whatman paper saturated respectively with: Tray 1: 0.2 M NaOH, 1.5 M NaCl; Tray 2: 0.4 M Tris-HCl, pH 7.6, 2 x SSC; Tray 3: 2 x SSC. The filter (plaque side up) was placed in tray 1 denaturing solution for 5 minutes at room temperature. Then it was transferred (plaque side up) to tray 2, neutralised for 5 minutes at room temperature and then transferred to tray 3 to rinse for 5 minutes. After being air-dried, the filters were fixed in a vacuum-drier at 80°C for two hours.

The filters were hybridised and washed as described in the probe preparation and hybridization protocol. The positive signals were confirmed by matching the signals from duplicate films. The positive plaques were those corresponding to the positive signals on the films. As the first screening was at high density it was impossible to get separate plaques so the areas where showed a positive signal was cored out into SM buffer and second or third screenings were carried out until separate positive plaques were obtained.

2.5 DNA sequencing, sequence analysis and primer designing

DNA sequencing was performed on an ABI 377 using dye primer automatic sequencing.

http://www.williamstone.com/primers/ was used to design primers either for PCR amplification or sequencing

http://www.ncbi.nlm.nih.gov/blast/ was used for searching for homologous PAL and ACC oxidase sequence in the gene bank database.

http://evolution.genetics.washington.edu/phylip/software.html/ or CLUSTAL W (1.8) multiple sequence alignment from http://workbench.sdsc.edu/CGI/BW.cgi was used for multiple alignment.

The GCG package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group), was used for comparing two sequences (Bestfit) and assembling sequences.

2.6 Plasmid constructs or vectors for promoter analysis

pKG: A plasmid with the GUS gene (*uidA* gene) and the nopaline synthase (nos) teminator inserted in pUC19 (Figure 2.1) (original name is polyKozakGUSTER, a gift from Andrew Goldsbrough, Long Ashton, UK). This plasmid had no promoter upstream of GUS, so the MePAL2 promoter could be fused to the 5' end of GUS to characterize the PAL promoter activity or be used as a negative control for transient GUS assay in this research.

p35S: (original name is pCamvGUSTER) was a derivative of pKG (originated from pUC19), with the CaMV 35S promoter fused in front of GUS in pKG (Figure 2.1) (a gift from Andrew Goldsbrough). This construct was used as a positive control for GUS assay.

p35H: (original name is pHX4) contained a CaMV35S-driven hygromycin resistant gene (*hph*) and the *nos* terminator (Figure 2.1). It was a construct from constructed from pUC19 by ILTAB and was used as selection marker for rice transformation (co-bombardment).

pMON977: a modified binary vector from ILTAB. It contains a CaMV 35S-driven kanamycin resistance gene (KAN) and nos terminator as selection marker in plants. The streptomycin resistance gene (*Str*) is the bacterial selection marker in this plasmid. In addition to those genes, there are cassava mosaic virus (CVMV) and the E9 3' terminator (a polyadenylation site in the pea rbcS-E9 gene (Li and Hunt, 1995) in this plasmid (Figure 2.1). This vector was used for making MePAL promoter-GUS constructs for analysis of the promoter by transient and stable assays in cassava and rice.

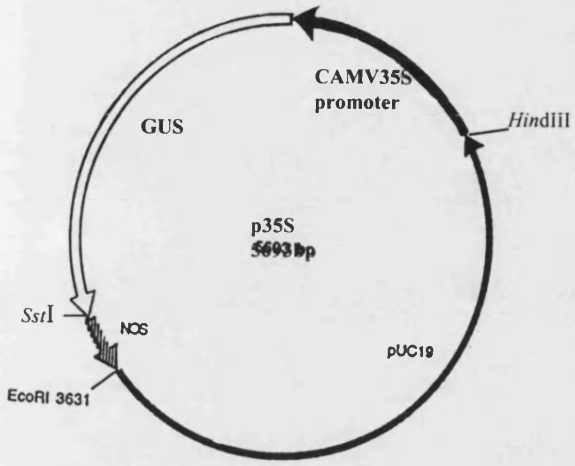
Fig 2.1 Maps of p35S, pKG, pMON977 and p35H

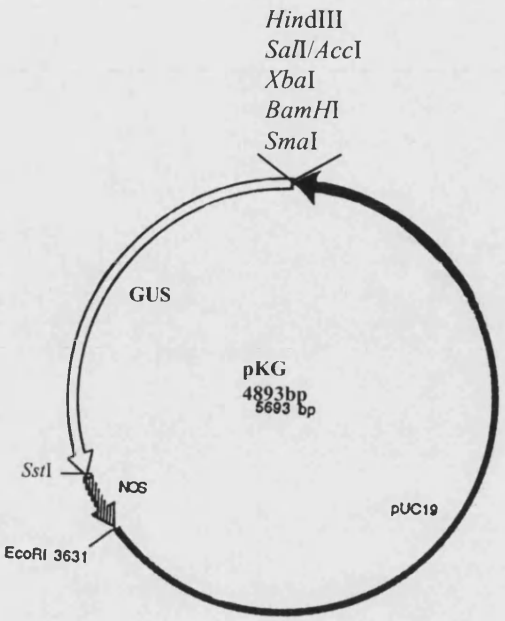
A: (Top left) p35S is the construct used as a positive control for transient assays. The GUS gene was driven by the cauliflower mosaic virus 35S promoter (CaMV35S promoter) and the nopaline synthase terminator *(nos)* was attached to the GUS gene.

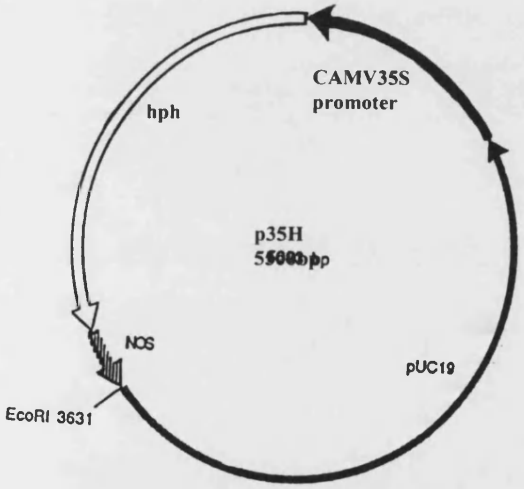
B: (Top right) pKG included a promoterless GUS gene and nos terminator. The construct was used as a negative control for transient assays.

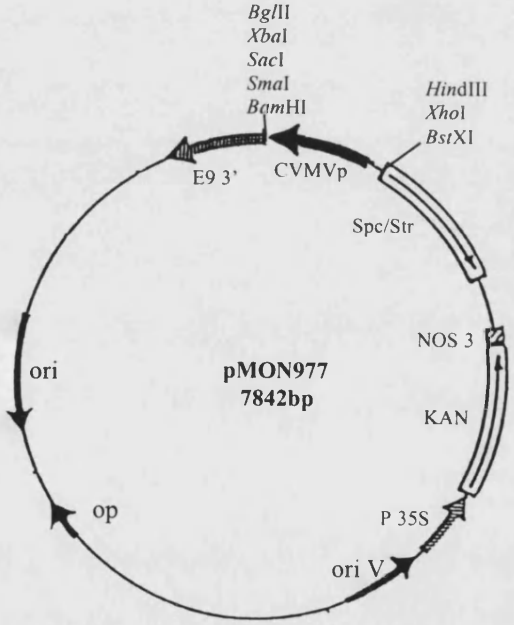
C: (Bottom Left) p35H included a CaMV35S-driven hygromycin resistant gene (*hph*) and nos terminator. The construct was used as a selection marker for rice co-bombardment.

D: (Bottom right) pMON977 contains a CaMV 35S driven Kanamycin resistant gene (KAN) and *nos* terminator as a plant selective marker gene. Streptomycin resistant gene (*Str*) is the bacterial selective marker in this plasmid. In addition, there are cassava mosaic virus (CVMV) and E9 3' terminators in this plasmid.


CAMV35S
promoter
GUS
HindIII
p35S
5693bp
SstI
NOS
pUC19
EcoRI 3631


HindIII
SalI/AccI
XbaI
BamHI
SmaI
GUS
pKG
4893bp
5693 bp
SstI
NOS
pUC19
EcoRI 3631


CAMV35S
promoter
hph
p35H
5693bp
NOS
pUC19
EcoRI 3631


BglII
XbaI
SacI
SmaI
BamHI
HindIII
XhoI
BstXI
E9 3'
CVMVp
Spc/Str
NOS 3
ori
pMON977
7842bp
KAN
op
P 35S
ori V

2.7 Cassava transformation and regeneration

2.7.1 Transient assays

Transient assays were carried out in embryogenic suspension cells and young leaves from *in vitro* plantlets.

Plant Material: Somatic embryogenesis was induced from leaf material of cassava cultivar TMS 60444 grown in glass jars on MS basal medium supplemented with 20 g/l sucrose. Immature leaf lobes were excised and placed on MS medium supplemented with 20 g/l of sucrose and 50 μM picloram. Secondary continuous embryogenesis was promoted by transferring the primary embryogenic tissue onto fresh medium of the same type. Friable embryogenic callus (FEC) was induced from the organized embryogenic tissue after being cultured on GD2 50P (Table 2.1). Embryogenic suspensions were initiated from FEC and cultured in 250 ml flasks containing 50 ml liquid SH6 50P (Table 2.1) (Taylor *et al.*, 1996). The suspensions were incubated in a shaker at 150 rpm at 25°C with a photo-period of 16 hours. The culture medium was replaced with fresh SH6 50P medium every 3 days. The suspensions were sub-cultured every three weeks.

Preparation of Suspension Cells for Transient Assay: Twelve to fourteen day old suspensions were drawn in and out of a 10 ml syringe several times in order to dissociate clumps of embryogenic tissue into smaller units. The resulting suspension was sieved, and the fraction composed of units in the range of 100-500 μm in diameter was resuspended in SH6 50P medium. Settled cell volume (SCV) was determined by transferring the tissue to a 15 ml graduated test tube to settle for 30 minutes. Aliquots of 250 μl SCV were pippeted onto plastic mesh and the tissue placed in the centre of Petri dishes and spread evenly over the surface to obtain a monolayer of embryogenic units. Prior to bombardment, the surface water of the tissue was removed by transferring the mesh with its tissue onto a double layer of dry filter papers. After bombardment, the filters were transferred onto solidified GD6 50P (Table 2.1) and cultured for 3 days at 25°C in a photoperiod of 16 hour/day, before the GUS assay was carried out to test the transient activity of the promoters.

Cassava plantlets, MCol 1505 were propagated *in vitro* at 26 °C, 16 hours light, in MS medium in a plant growth chamber. New leaflets (folded leaves or not fully open) were excised and placed onto MS2 medium (Table 2.1) 2 hours before they were used for transient assays.

Microparticle Bombardment: All bombardments were conducted with the Biolistic PDC-1000/HE system (BIO-RAD, Hercules, California). Three milligrams of 1.0 μm diameter gold particles was washed once with 100% ethanol, twice with sterile distilled water and resuspended in 50 μl water in a siliconized 1.5 ml of centrifuge tube. Five micrograms of plasmid DNA, 20 μl spermidine (0.1 M) and 50 μl calcium chloride (2.5 M) were added to the gold suspension under continuous vortexing. The mixture was incubated at room temperature for 10 minutes and pelleted at 1000 rpm for 10 seconds, the pellet was resuspended in 60 μl cold 100% ethanol. Eight to nine μl of resuspended gold particles were distributed onto each macrocarrier. Tissue samples were placed 8 cm beneath the stopping plate of the gun and bombarded at 1100 psi and 27 in of Hg vacuum.

Histological GUS Assay: Three days after bombardment, bombarded tissue was tested for GUS activity. GUS assay buffer (modified from Jefferson, 1987) (Section 2.9) was pippeted carefully onto the bombarded tissue so as not to disturb it. After 2 hours incubation at 37°C, the tissue was gently washed twice with 70% ethanol. In order to maintain the position of suspension cells on the plate, the plate that contained the tissue and the GUS assay buffer was put at an angle. Then the tissue was slowly rinsed by adding the ethanol at the top of the plate and collecting it at the bottom. The blue spots were counted under the microscope.

2.7.2 Stable transformation

Three to six month old TMS 60444 embryogenic callus was used as the target tissue for stable transformation. Cell clusters from embryonic callus of TMS60444 were disaggregated by pumping in and out of a 10 ml syringe and sieved to isolate

embryogenic units 100-500 μM in diameter. The rest of the procedure was as for transient assay particle bombardment.

After bombardment, the tissue was transferred by washing with 15 ml SH6 50P liquid medium to a small bottle in laminar flow-hood to avoid contamination and the tissue was cultured on a shaker at 150 rpm at 25°C with a photo-period of 16 hours for three days in order to permit recovery from bombardment. Then the medium was replaced with selection medium, SH6 50P medium supplemented with 25 μM paromomycin (Table 2.1). After four days in liquid selection medium, the embryogenic tissues were plated onto solidified selection medium (solidified GD6 50P medium supplemented with 25 μM paromomycin) (Table 2.1). Ten days later, and every week thereafter, surviving pieces of yellowish tissue were transferred individually onto fresh solidified selection medium for three weeks. After another round of three weeks in selection medium to avoid escape of non-transgenic callus, tissues that still grew and established colonies of FEC were then considered putative callus resistant lines or callus lines. These lines could be used for GUS assays, PCR analysis for transgenes and part of the callus from each line could be transferred to the regeneration steps to get the transgenic plant.

Regeneration of plantlets from callus involved three stages including differentiation of embyo, shoot and root induction. For each stage, different media based on MS basal medium (Murashige and Skoog, 1962) (supplemented with 20 g/l sucrose) with varied concentrations of α-naphthalene acetic acid (NAA) or 6-benzylaminopurine (BAP) were used. To obtain cotyledons, transgenic embryogenic calli first were transferred to MS2 5NAA (Table 2.1). In this stage, a few of lines can reach cotyledon stage while most of the lines were still at the globular or torpedo embryo stage. Those lines at globular or torpedo embryo stage were transferred to fresh medium with reduced NAA, which was MS2 1NAA, MS2 0.1NAA or MS2 to induce the cotyledons. Once cotyledons were induced, they were cultured in MS basal medium supplemented with 20 g/l sucrose and 2 μM of BAP (MS2 2BAP) to induce shoot. Roots were induced in MS basal medium supplemented with 20 g/l of sucrose (MS2) after a first a real leaf appeared in MS2 2BAP medium. The sequence of transfers to various media during the

course of regeneration is described in Figure 2.2: Production of Transgenic Cassava Plants.

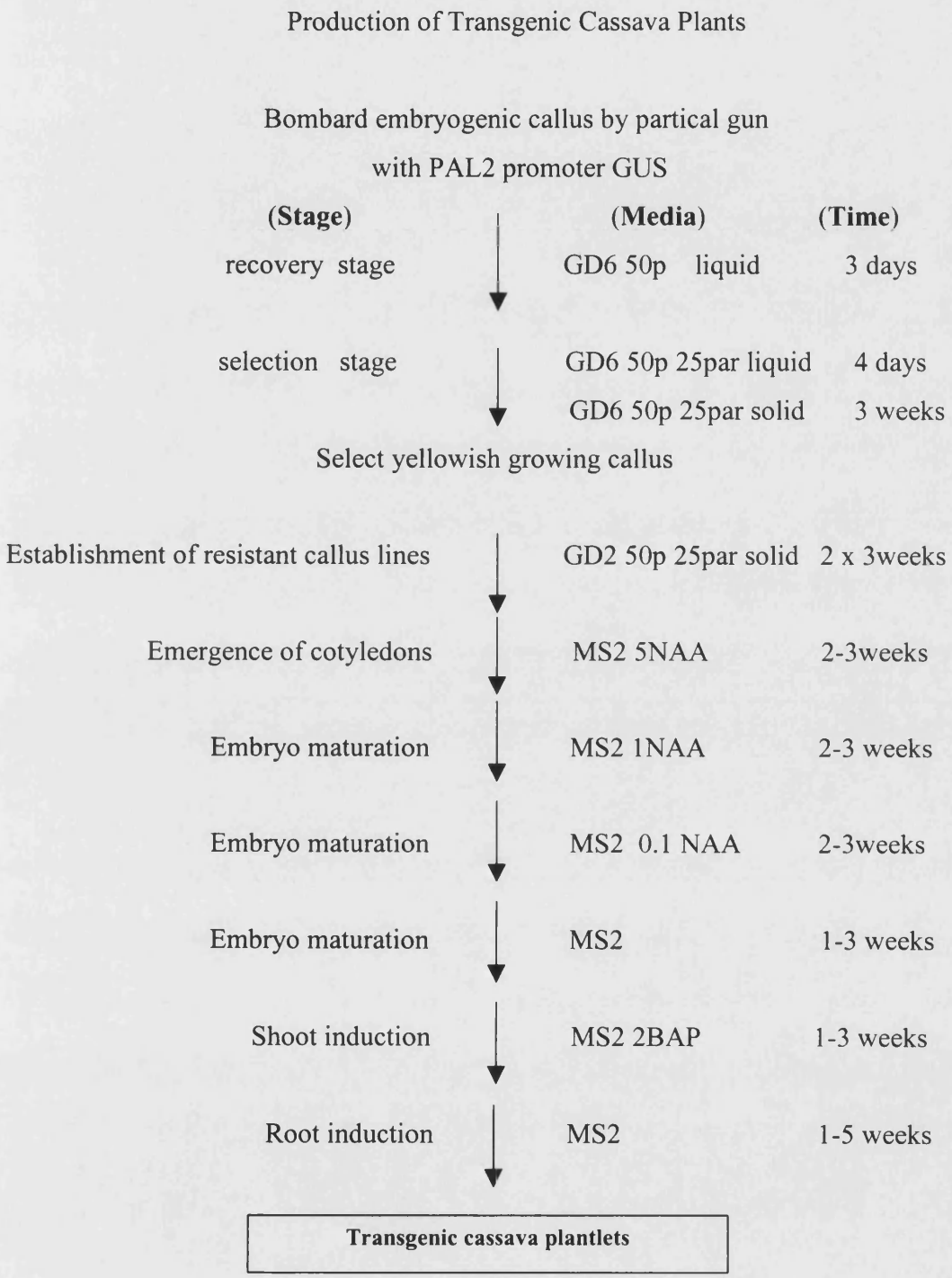


Figure 2.2 Flow chart of cassava transformation

(Note: The media are shown in Table 2.1)

Table 2.1 cassava transformation and regeneration media (per litre)

Name of medium	Description	Function
SH6 50P	60 g sucrose, 1 x SH[1] mixture, 50 μM picloram , pH 5.7	Liquid medium for friable embryogenic suspension cell
GD6 50P	60 g sucrose, 1 x GD[2] mixture, 50 μM picloram , pH 5.7,	Solid medium for maintaining embryogenic callus
GD2 50P	20 g sucrose, 1 x GD mixture, 50 μM picloram	Solid medium for growing embryogenic callus
GD2 50P 25 par	Add 25 μM of paromomycin to GD2 50P	Solid medium for selecting transgenic callus
MS2	20 g sucrose, 1 x vitamin[4], MS[3] mixture 4.31 g, pH5.7	Solid medium for growing *in vitro* cassava or root induction
MS2 5NAA[5]	Add 5 μM of NAA to MS2 medium	Solid medium for differentiation of embryos
MS2 1NAA	Add 1 μM of NAA to MS2 medium	As above
MS2 0.1NAA	Add 0.1 μM of NAA to MS2 medium	As above
MS2 2BAP[6]	Add 2 μM of BAP to MS2 medium	Solid medium for shoot induction

1. SH: Schenk and Hildebrandt basal medium (Sigma)
2. GD: Gresshoff and Doy medium (Duchefa)
3. MS: Murashige and Skoog (1962) (GibcoBRL)
4. MS vitamin (Sigma)
5. NAA: α-naphthalene acetic acid (Sigma)
6. BAP: 6-benzylaminopurine (Sigma)

All the media used for cassava transformation were solidified by the addition of 7.8 g/l Agar Noble (DIFCO laboratories) and adjusted to pH 5.7 prior to autoclaving.

2.8 Rice transformation

Oryza Sativa L. Japonica cv. Taipei 309 (TP309) was used for rice transformation. The rice transformation system developed by Chen *et al.* (1998) was used in this research.

2.8.1 Embryogenic callus induction from rice seeds and regeneration after gene transformation

Embryogenic rice callus was used for the target tissue in the rice transformation. Mature rice seeds were surface-sterilized in 70% ethanol for 1 min, followed by 2.6% (w/v) sodium hypochlorite for 45 minutes and then were rinsed three times with sterile distilled water. The surface-sterilized seeds were plated on Petri dishes containing solidified NB medium (Section 2.8.2) to induce embryogenic callus. After two to three weeks of culturing in the dark at 25-26°C, primary callus, approximately 1 mm in diameter, was removed from the scutellum to new NB medium. Embryogenic callus, 0.5-1 mm in diameter, was amplified in the new NB medium and small, compact new calli was subcultured every 2 weeks in fresh NB medium until used for gene transformation. Callus tissues, which were not older than three months, were subjected to particle bombardment 9-14 days after the last subculture. Embryogenic calli, 1 to 3 mm in diameter, were chosen and placed on NBO osmotic medium (Section 2.8.2) to form a circle about 2.5 cm across in a Petri dish and cultured for four hours prior to bombardment.

Sixteen to twenty hours after bombardment, tissues were transferred from NBO medium onto NBH40 hygromycin selection medium (Section 2.8.2), ensuring that the bombarded surface was facing upward, and incubated in the dark at 25-26°C for two weeks. Newly formed callus was then separated from the original bombarded explants and placed nearby on the some medium. Following an additional two weeks, a few relatively compact and opaque calli were visually identified, and transferred to PRH50 pre-regeneration medium (Section 2.8.2) for one week in the dark at 25-26°C. Those identified calli were considered as hygromycin resistant lines. Calli became more compact and opaque 1 week after transferred onto PRH50 medium. These calli were then subcultured onto regeneration medium for two weeks under a 16 hours

photoperiod per day at 25-26°C. During this period the shoots were induced and regenerating shoots were then transferred to Magenta boxes containing ½ MSH50 medium (Section 2.8.2) to induce the roots. Once plantlets had reached the top of the Magenta boxes, they were transferred to ½MSH50 liquid medium for a week under the same condition (16 hours photoperoid per day at 25-26°C). After a week in the liquid medium, the rice plantlets were moved to 13 cm pots in a greenhouse with a 14 hours light period at 30°C and in the dark at 21°C.

2.8.2 Media for rice transformation

Stock solution:

1. N6 Macro-elements 20x

KNO_3	56.6 g/l	KH_2PO_4	8.0 g/l
$(NH_4)_2SO_4$	9.26 g/l	$CaCl_2.2H_2O$	3.3 g/l
$MgSO_4$	1.8 g/l		

2. B5 Micro-elements 100x

$MnSO_4.H_2O$	785 mg/l	$ZnSO_4$	112 mg/l
KI	75 mg/l	$Na_2MoO_4.2H_2O$	25 mg/l
H_3BO_3	300 mg/l	$CuSO_4.5H_2O$	2.5 mg/l
$CoCl_2.6H_2O$	2.5 mg/l		

3. Fe-EDTA (100x)

   Ethylene-Diamine-Tetraacetic acid, Ferric-Sodium Salt 4.15 g/l (Sigma E-6760)

4. B5 Vitamin (100x): Gamborg's vitamin powder 2.8 g/ 250 ml (Sigma G-2519)

Medium for rice transformation unit: per litre

NB: 50 ml of N6, 10 ml of B5, 10 ml of B5 Vitamin, 10 ml of Fe-EDTA, Sucrose 30 g, Proline 500 mg, Glutamine 500 mg Casein Enzymatic Hydrolysate 300 mg, 2.4-D 2 mg,
adjust pH 5.7 then add 2.6 g phytagel per liter for solidification

NBO: Add 47 g of mannitol and 47 g of sorbitol to NB medium before autoclaving

NH30: Add 30 mg of hygromycin to NB medium after autoclaving

NH40: Add 40 mg of hygromycin to NB medium after autoclaving

PRh50: Add sterilized 2 mg of BAP, 1 mg of NAA and 5 mg of ABA to NB medium without 2,4-D after autoclaving

RNH50: Add sterilized 3 mg of BAP and 0.5 mg of NAA to NB medium with out 2,4-D after autoclaving

1/2mSH50: ½ MS* salts, ½ B5 vitamin (see stock solution), 10 g sucrose, and add 50 mg hygromycin after autoclaving

Note: * MS: Murashige and Skoog (1962) (GibcoBRL)

2.8.3 Particle bombardment

Plasmid p35H containing the hygromycin resistant gene (*hph*) as a selection marker, was co-bombarded with the constructs of MePAL2 promoter/GUS fusions. The p35H plasmid and a MePAL2 promoter-GUS plasmid were mixed at a 1:7 molar ratio and made up to total of 5 μg of DNA for one set of experiment (Section 2.7.1).

The rice transformation was carried on the Biolistic PDC-1000/HE system (BIO-RAD, Hercules, California) under the same conditions as cassava transformation. Gold preparation and bombardment procedures for the rice calli were the same as cassava transformation (section 2.7.1) except that DNA coating was a mixture of two plasmids (total 5 μg), instead of 5 μg DNA of a single plasmid.

2.9 Analysis of transgenic cassava and rice

2.9.1 Localisation of GUS expression

Histochemical analysis of transgenic cassava and rice for GUS localisation was performed at different plant developmental stages using methods as described by Jefferson *et al.* (1987). Transgenic cassava and rice tissues such as stems, roots, leaves, leaf sheath, petiole and tuber were free-hand sectioned and then stained with X-gluc (5-bromo-4-choloro-3-indol-l-glucuronide) GUS assay buffer for 2 hours at 37°C in the dark. Assay buffer contained potassium ferrocyanide 8 mM, potassium ferricyanide 8 mM, X-glu: 5 mg/10ml, Triton X-100 0.2%, sodium phosphate buffer 0.1 mM, pH 7.0.

After staining with GUS assay buffer, chlorophyll was removed by washing three times with 70% ethanol, and sections were mounted for observation under a microscope.

.

2.9.2 Quantification of GUS activity

About 1 g – 2 g of plant tissue such as leaf, root, stem or petiole were grounded in liquid nitrogen to a fine powder. The powder was put into a 1.5 ml micro-tube containing 500 µl of GUS extraction buffer (50 mM $NaPO_4$, pH 7.0; 10 mM beta-mercaptoethanol; 10 mM Na_2EDTA; 0.1% (w/v) Sodium lauryl Sarcosine; 0.1% (v/v) Triton X-100) (Jefferson, 1989). The mixture was vortexed vigorously and centrifuged at 15,000 rpm, for 20 minutes at 4°C. The supernatant was filtered through miracloth into another centrifuge tube. The extract was tested immediately or stored at -70°C.

Total protein concentration was measured using the Bradford Protein Assay Kit (BioRad, Mississauga, Ontario) at 955 nm in microtiter plates. Bovine serum albumin (1 µg, 2 µg, 3 µg, 4 µg, 5 µg) was used to make a standard curve to determine the protein concentration.

An aliquot of 0.5 ml MU assay buffer (1 mM 4-Methyl umbelliferyl beta-D-glucuronide in GUS extraction buffer) in a 1.5 ml micro-tube was pre-warmed at 37°C. Then 100 µl of extract was added to the micro-tuber, mixed thoroughly by vortexing and incubated at 37°C. At every 20 minutes, 100 µl of the mixture was transferred to a 1.5 ml micro-tube containing 900 µl of stop buffer (0.2 M Na_2CO_3) at room temperature. Methylumbelliferone (MU) concentrations of each sample were determined with a spectrofluorimeter (Perkin-Elmer LS), at an excitation of 365 nm and an emission of 455 nm. Relative fluorescence was read and calibrated using a MU standard curve, which was made with MU 100 µg, 200 µg, 300 µg, 400 µg, and 500 µg. GUS activity was expressed as the concentration of 4-MU released per minute per mg protein.

CHAPTER THREE: ISOLATION OF MEPAL GENOMIC CLONES

3.1 Introduction

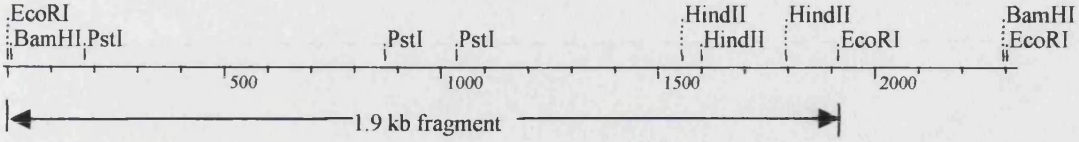
The potential importance of PAL genes in the development of PPD in cassava has been recognised and effort has been put into the isolation of PAL genes. Using degenerate primers from conserved regions of PAL genes in other plant systems, a cassava PAL gene fragment (MePAL) was amplified from genomic DNA (Pereira *et al.*, 1999). From a PPD-related cDNA library, two PAL clones cMePAL1 and cMePAL3 were isolated using bean *Phaseolus vulgaris* gPAL2 DNA as a probe (Han, 2000). Limited expression data, especially spatial expression, on these genes were obtained. To fully understand the role of PAL genes in cassava, full length sequences of these genes would be required, in order to analyse the regulation of the PAL genes and their temporal and spatial expression patterns using promoter-GUS fusions in transgenic plants. This chapter reports the isolation of cassava genomic PAL clones by screening a cassava genomic library and by amplifying genomic fragments using PCR.

3.2 Organization of cassava PAL genes

It has been shown in many plant species that PAL is encoded by a multi-gene family, with the number of genes ranging from two to six in most cases. The number of PAL genes isolated from cassava demonstrate that PAL is also encoded by a multigene family in this crop. To gain more information about the size of the PAL gene family, which would help to estimate the scale of the screening of the genomic library, cassava genomic DNA was digested with restriction enzymes and Southern hybridisation was carried out.

The genomic DNA was digested separately with the restriction enzymes, *Eco*RI, *Xba*I, *Hin*dIII and *Bgl*II. The digestions were then run in a 0.8% agarose gel and the gel was blotted on to Hybond-N^+. The probe used for hybridization was made from the 1.9 kb *Eco*RI fragment of the MePAL1 cDNA (Figure 3.1) covering 85% of MePAL1 coding region.

From Figure 3.1, it can be seen that multiple bands were present in all lanes, in which cassava genomic DNA was either digested with restriction enzymes that do not cut the region covered by the probe (*Eco*RI, *Bgl*II and *Xba*I) or digested with the restriction enzyme that cuts the probe region (*Hin*dIII). When genomic DNA was cut by *Eco*RI, seven bands were observed using the cMePAL1 1.9 kb probe. These bands ranged from one to 23 kb and five of these bands were weak and the other two were strong. In lane *Xba*I, one broad strong band (17 kb) and one weak band (8 kb). In lane *Hin*dIII, there were six weak bands and three strong bands, ranging from 0.8 kb to 6.5 kb. In the *Bgl*II lane, four weak and two strong bands were observed, which ranged from about 1.5 to 23 kb. These data indicate that PAL was encoded by a multigene family.



A: Map of MePAL1 cDNA(2303bp) and the location of 1.9 kb probe

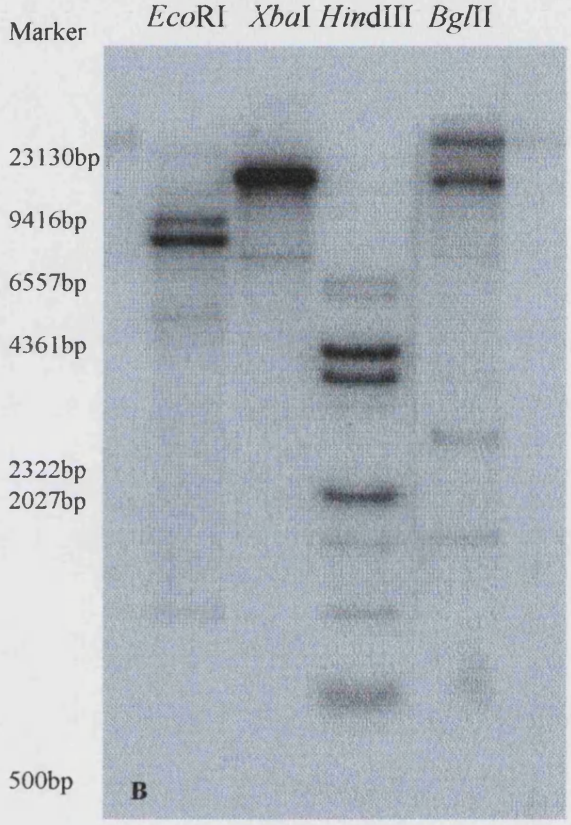


Figure 3.1 Southern analysis of cassava PAL genes.
A: Restriction map of MePAL1 cDNA and the location of 1.9 bp *Eco*RI fragment that was used as probe for Southern hybridisation.
B: Cassava genomic DNA (10 µg) extracted from young leaves was digested respectively with restriction enzymes *Eco*RI, *Xba*I, *Hin*dIII and *Bgl*II. The digestions were then run in 1 x TBE 0.8% agarose gel and subsequently blotted against Hybond-N^+. Hybridization was carried out at 60°C overnight as described in section 2.3.5 (chapter 2). After hybridization, the membranes were washed finally at low stringency (60°C, 1 x SSC 0.1% SDS, 2 x 10 min). The membrane was exposed to film between two intensifying screens overnight.

3.3 Isolation of genomic MePAL2 clones from a cassava genomic library

3.3.1 Titering, screening cassava genomic library

Before screening the cassava genomic library, it was titered first. A duplicated dilution series of the cassava genomic library indicated that the titer of the phage was 2.1 x 10^8 pfu/ml. In order to isolate a clone containing PAL DNA with 99% probability, 2.1 x 10^6 (10 µl) plaques (see Section 2.4 screening genomic library) were plated out on 4 plates (90 mm in diameter) and screened for PAL clones using the probe made from the 1.9 kb *Eco*RI fragment of MePAL1 cDNA. Hybridization was performed at 60°C overnight and the membranes were washed with 0.1 x SSC and 0.1% SDS at 60°C. Four positive plaques were isolated after the first screening. In order to get separate positive plaques a secondary screening was performed under the same condition but the density was much lower than the first screening at about 200-300 plaques on each plate. For each positive, one plate was screened, and one positive plaque was selected from each plate and designated as 1A, 2A, 3E and 4A. One example of the second screening is shown in Figure 3.2.

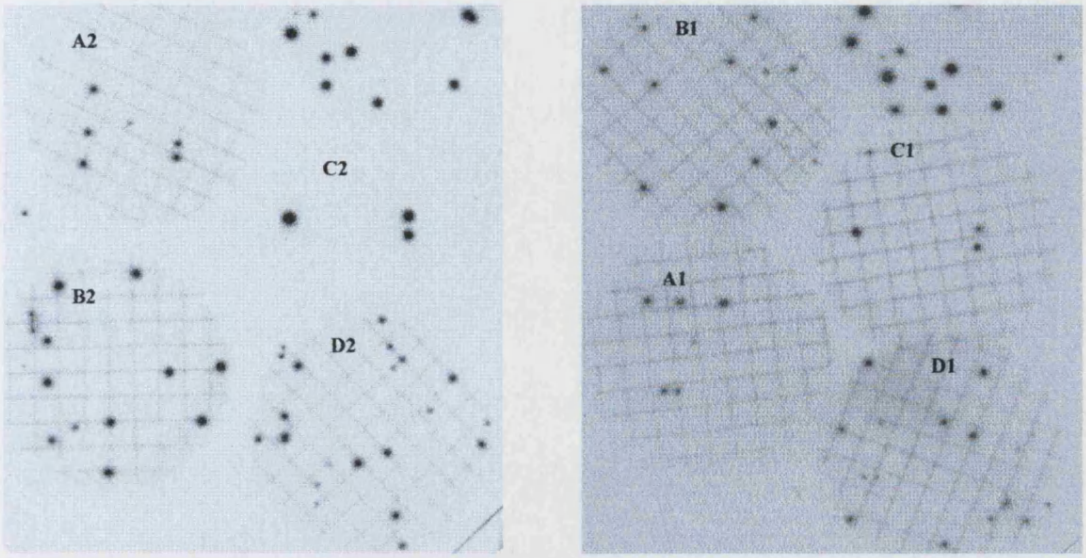


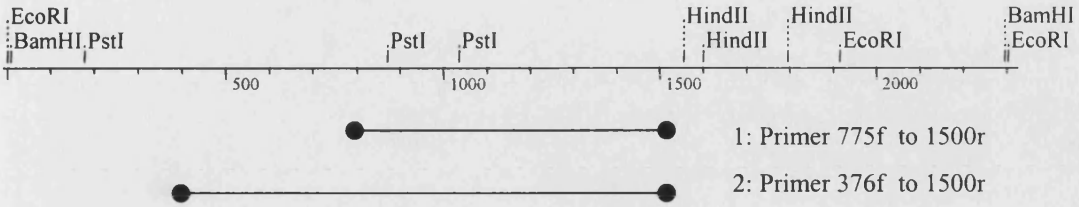
Figure 3.2 Secondary screening of the cassava genomic DNA library for PAL clones.

Plaques were lifted onto Hybond N+ membrane from four plates representing four positives from the first screening. Duplicate plaque-lifts were also done for each plate. Membranes were then denatured and neutralised as described in Section 2.3.6 The membranes were hybridised with the MePAL1 1.9kb *Eco*RI probe at 60°C overnight and washed with 0.1 x SSC / 0.1% SDS for 2 x 10 min. The membranes were exposed to film and developed. A1/A2, B1/B2, C1/C2 and D1/D2 are duplicates from four plates. The matched signals from duplicated membranes were positives, such as the five spots in A1 and A2, and the three spots in C1 and C2.

3.3.2 Confirmation of the identity of cassava PAL genomic clones

To test the putative positive clones or plaques from the genomic library and check if they were genomic PAL clones, primers, which were originally designed from MePAL1 cDNA for sequencing, were used to amplify relevant regions of the lambda DNA of the positive clones. Two forward and one reverse primers were used: 376f (5'CAAGGCCAG TAGTGACTGGG), 775f (5'AGTCGGTTAGGGCCCAATGG) and 1500r (5'GCCAT GGCAATTTCAGCTCC). Their locations in MePAL1 cDNA are shown in Figure 3.3A. The PCR products obtained using these two sets of primers with lambda clones as templates are shown in Figure 3.3B & C.

Firstly, it can be concluded based on the PCR products with primers 775f / 1500r (Figure 3.3B), that all the four putative positive clones were PAL genes, as from all of them DNA bands were amplified, the sizes of which were the same as the control which was amplified from the MePAL1 clone. Secondly, there must be an intron in the region covered by primers 376f and 1500r since the PCR products from the four genomic clones were about 800 base pairs, larger than the PCR product from the PAL cDNA (Figure 3.3C). Furthermore the location of the intron could be narrowed down to the region between 376f and 775f, as there was no intron between 775f and 1500r. These results also indicate that these four clones may be identical in the regions covered by primers 775f and 1500r, at least the size of these regions are the same even if their sequences may be different in these regions. The PCR product from 1-A clone with 376f and 1500r primers was purified and sequenced. A Blast search with the partial 1-A sequence showed high similarity to the sequence of PAL genes in other plant species. A comparison of this sequence with that of MePAL1 using GCG (program shown in Section 2.5) showed that this sequence was not the same as MePAL1 and the similarity between these two sequences was 79% over about 400 bp.



A: Map of MePAL1 cDNA and location of the primers used for the PCR confirmation .

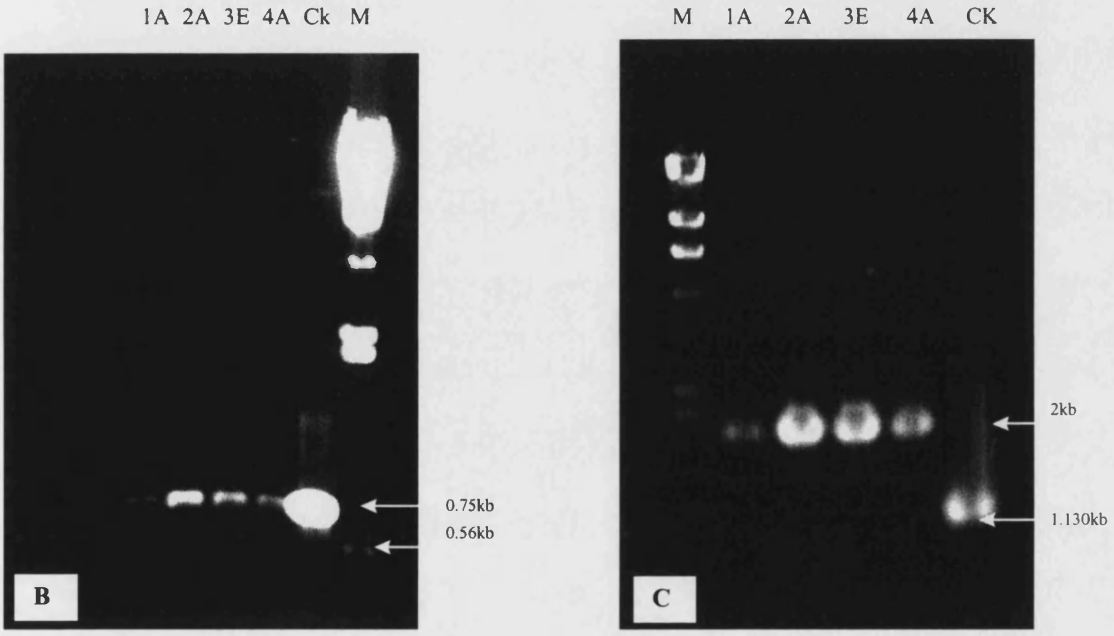


Figure 3.3 PCR test of the putative positive PAL clones isolated from the cassava genomic library.

A: Map of MePAL1 and location of the primers used for the PCR confirmation.
B: PCR using primers 775f and 1500r.
C: PCR using primers 376f and 1500r.
PCR was carried out with one cycle of 94°C 3 min., followed by 35 cycles of 94°C 1 min, 53°C 1 min, 72°C 3 min., and completed with 72°C for 10 min. The PCR products were run in 0.8% agarose (0.5 x TBE).
M: *HindIII* cut λ DNA marker. 1A, 2A, 3E and 4A are positive PAL clones. CK: PCR products from cDNA MePAL as a positive control.

3.4 Sub-cloning and sequencing a MePAL genomic clone

3.4.1 Restriction digestion and hybridisation

In order to identify appropriate fragment(s) for sub-cloning and subsequent sequencing, λ1A, 2A, 3E and 4A were single-digested with different restriction enzymes: *Eco*RI, *Hin*dIII and *Bgl*II, and run on an agarose gel (Figure 3.4B). The digested DNA was Southern-blotted to a nylon membrane, which was then hybridized separately with probes made from the 1.9 kb and 0.4 kb *Eco*RI cut fragments of MePAL1 cDNA (Figure 3.4A shows the location of the probes in MePAL1). The results of Southern hybridisation are shown in Figure 3.4C (using 1.9 kb cDNA MePAL1 probe) and Figure 3.4D (0.4 kb cDNA MePAL1 probe). The Southern blots showed that two bands, a 6 kb and a 2.6 kb released by *Hind*III digestion from 1A, 2A and 4A hybridised to the 0.4 kb and 1.9 kb MePAL1 probes, respectively. A 2.6 kb fragment released by *Hin*dIII from 3E hybridised to both probes. Two fragments, released from all 4 clones in the *Bgl*II digestion, of about 4 kb and 1.8 kb, hybridised with the 1.9 kb MePAL1 probe. The sizes of these four fragments, from 1.8 kb to 6 kb were suitable for sub-cloning. The *Eco*RI digestion released two bands, both nearly 10 kb, which were too large for subcloning. Figure 3.4, suggests either that there may be two different PAL family members amongst these four clones, as the four clones show two different digestion patterns, or maybe these four clones were from the same gene but contain regions of different sizes.

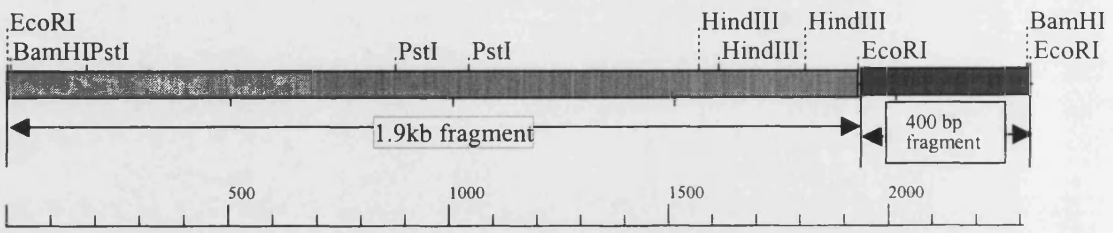


Figure 3.4A: Map of cDNA MePAL1(2303 bp) and the probes,1.9 kb and 0.4 kp *Eco*RI fragments that used for hybridization

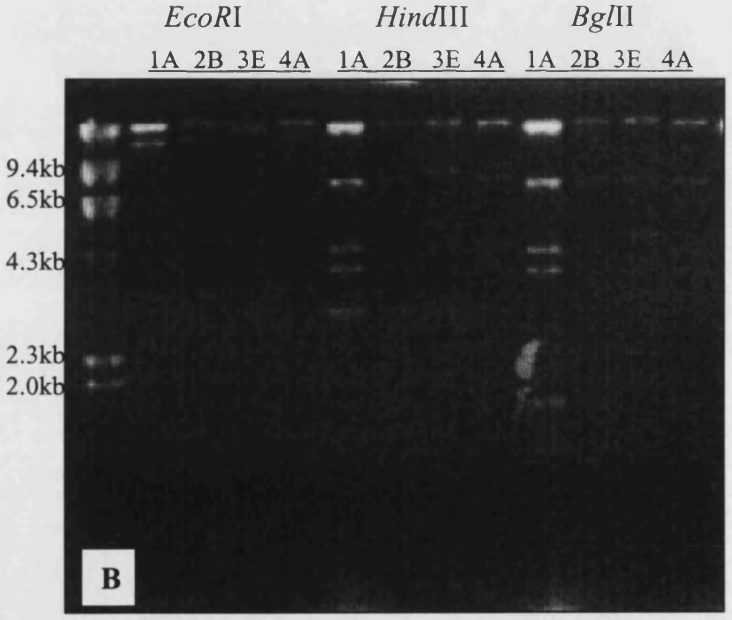


Figure 3.4B: Cassava genomic clone 1A, 2B, 3E and 4A were digested respectively with restriction enzymes *Eco*RI, *Hin*dIII and *Bg*lII. The digestions were then run in 1 x TBE 0.8% agarose gel and subsequently blotted against Hybond-N^+.

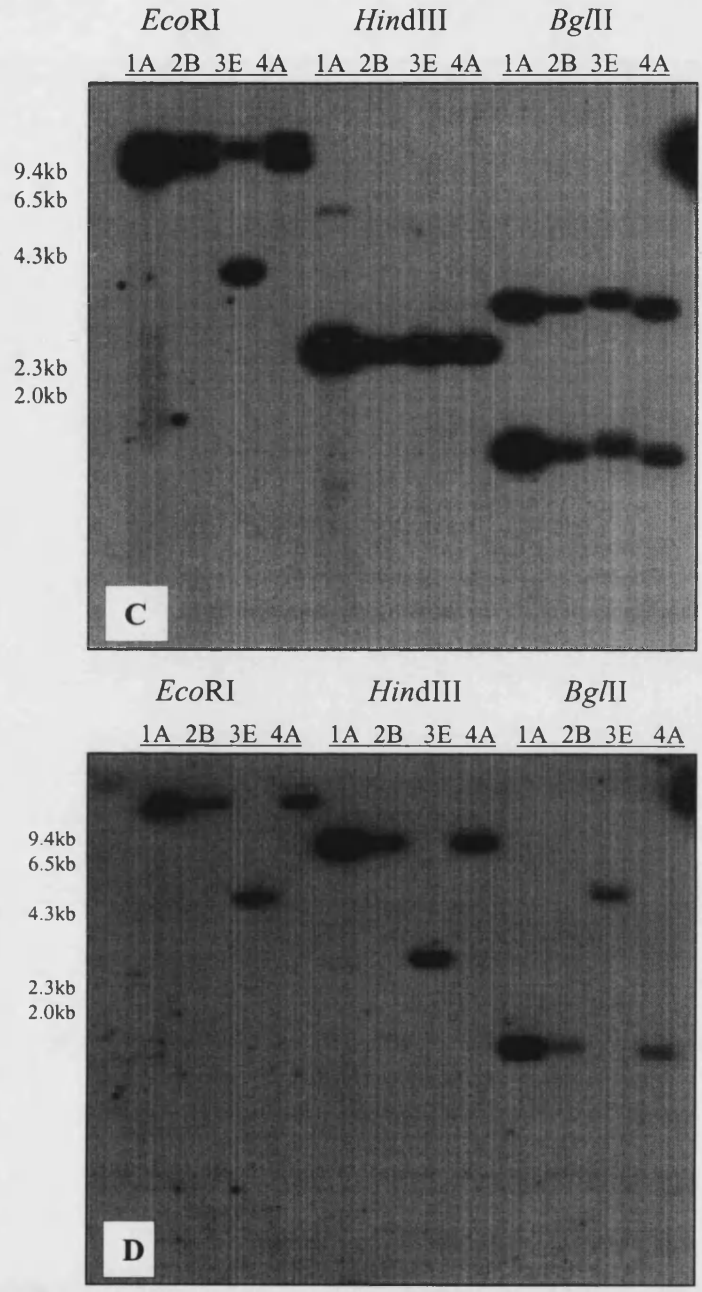


Figure 3.4 Restriction digestion of lambda MePAL clones and Southern hybridization.

A: The location of the probes (1.9 kb and 0.4 kb fragment) used for hybridization

B: Digestion of four genomic MePAL clones.

C: MePAL genomic clones hybridised with the1.9 kb fragment of MePAL1 cDNA.

D: MePAL genomic clones hybridised with the 0.4 kb fragment of MePAL1 cDNA.

Hybridization was carried out at 60°C overnight as described in section 2.3.6 (chapter 2). After hybridization, the membranes were washed finally at low stringency (60°C, 1 x SSC 0.01% SDS, 2 x 10 min). The membrane was exposed to film between two intensifying screens overnight.

3.4.2 Subcloning two *Hin*dIII fragments and sequencing

Lambda clone 1A was digested with *Hin*dIII. The resulting 2.6 kb and 6 kb fragments were purified and subcloned into the *Hin*dIII site of the plasmid pUC19, these were designated respectively as p1A26 and p1A60. The inserts were sequenced using universal primers first. The sequence of the 3' region of p1A26 showed 78.47% similarity with MePAL1 (BESTFIT of GCG package), while p1A26 5' sequence showed no similarity to PAL sequences from cassava or PALs from other known plant species. The p1A60 5'- sequence corresponded to the cDNA MePAL from 1427 base pair to 2025 base pairs with similarity 79.0%. While the p1A60 3' region was identified as part of the cloning vector lambda EMBL3 (100% identity in 627 bp overlap). It became clear that p1A26 contained the 5' flanking region and p1A60 contained the 3' flanking region of a cassava PAL gene, and that this gene was different from MePAL1. Therefore, this gene was designated as MePAL2.

The inserts of p1A26 were sequenced by designing primers from both directions based on sequenced regions. Since the 3' part of p1A60 corresponded to lambda EMBL3, p1A60 was sequenced to about 1000 base pairs from the 5' end. The sequences of each fragment was edited and assembled with the Gelassemble program in the GCG package.

In order to join the sequences of these two clones together, a forward primer from the end of p1A26 and a reverse primer from the beginning of p1A60 were used to amplify the relevant region from 1A phage clone. This PCR product was purified and sequenced. The sequence included parts of 1A26 and 1A60. The PCR fragment sequence enabled the sequences of 1A26 and 1A60 to be joined and assembled to give of final sequence of 4791 bp. The sequence strategy is shown in Figure 3.5 and the final sequence shown in Figure 3.6.

3.5 Structure of MePAL2

Based on comparison between the sequences of cDNA MePAL1, MePAL2 and other well studied plant PAL genes, together with the Open Reading Frame (ORF) analysis

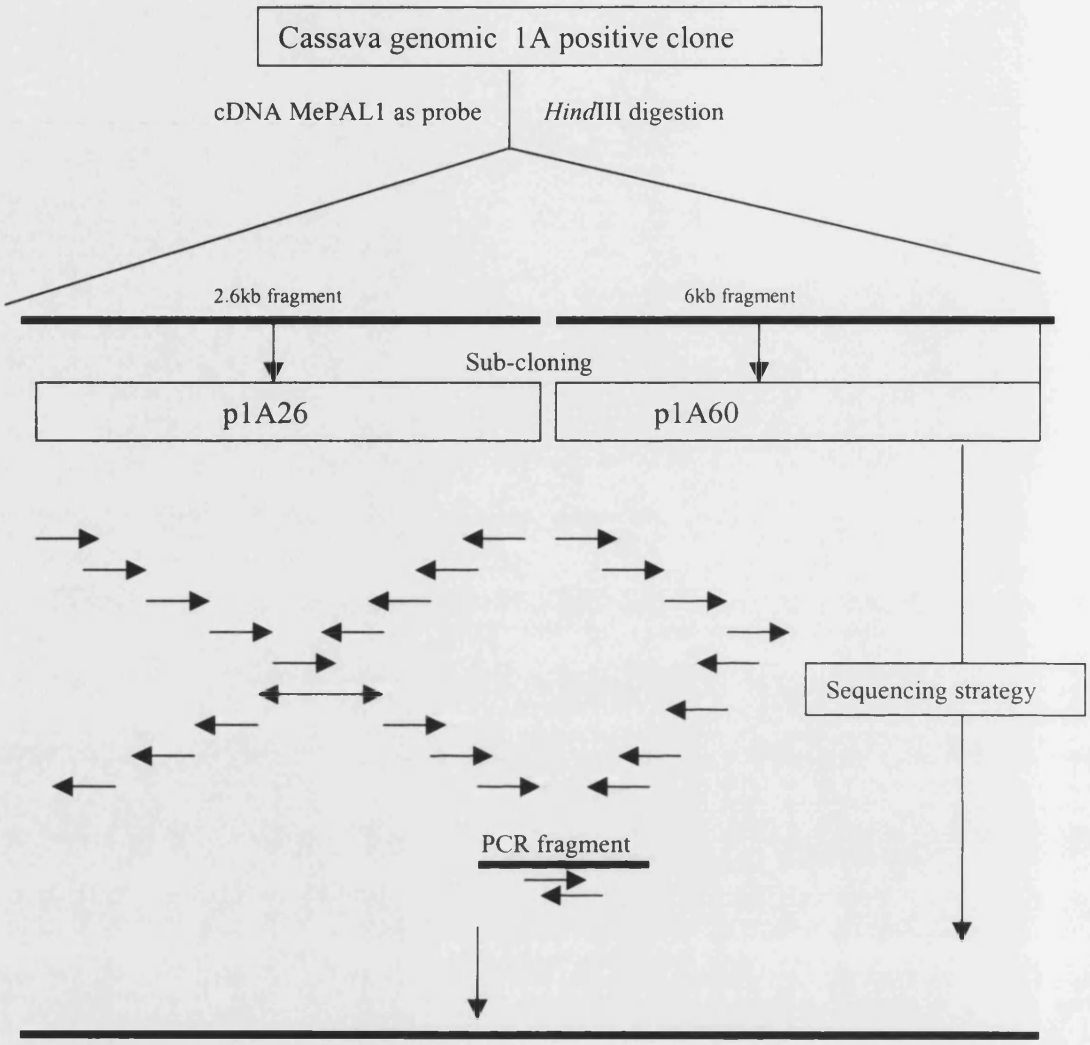


Figure 3.5a Schematic map of subcloning procedures and sequencing strategy of MePAL2 from the 1A genomic clone.

Genomic DNA of p1A was digested with *Hin*dIII and the resulted two *Hin*dIII fragments were subcloned into pUC19. The *Hin*dIII inserts were sequenced. The DNA region between 1A26 and 1A60 was amplified using PCR with primers designed from 1A26 and 1A60.

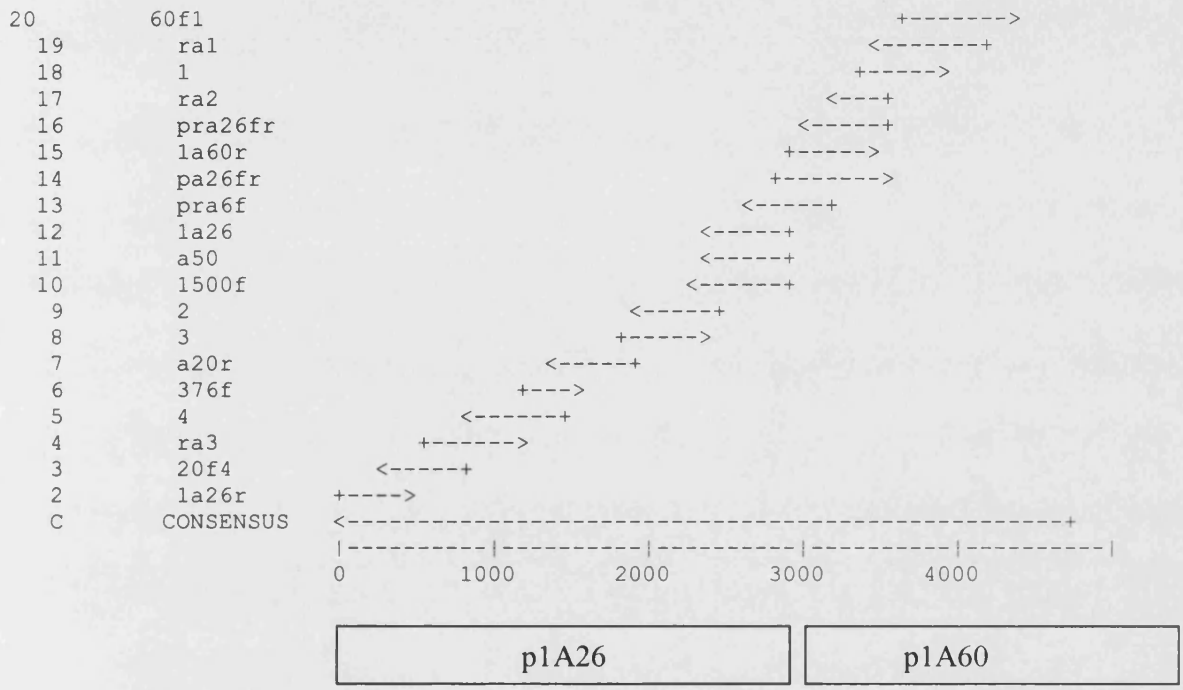
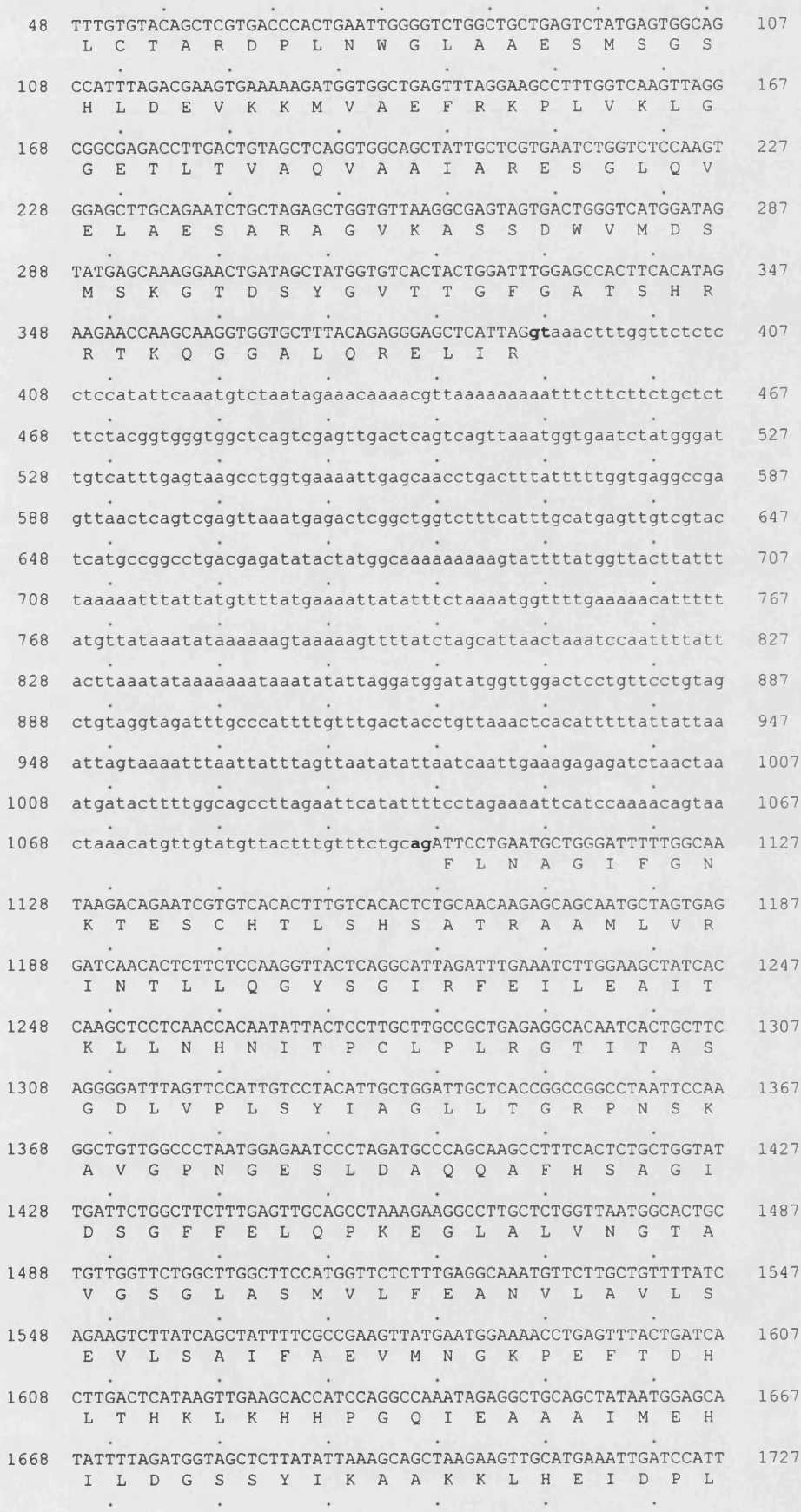


Figure 3.5b Schematic map of sequence assembly of all pieces of sequences from cassava PAL positive clone p1A by GCG program (section 2.5).

On the left is the number of each sequence that was determined by the primer on its right. The dashed line and arrow indicate the position of the determined sequence in relation to the assembled sequence and the sequencing direction. The scale of the sequence is indicated below the consensus at base-pair.

by the GCG package, the features of MePAL2 were identified (see Figure 3.6 MePAL2 genomic DNA sequence). It consisted of a 913 bp promoter region, two exons (388 bp and 1748 bp), one 711 bp intron and a 860 bp 3' flanking region. The predicted MePAL polypeptide was 712 amino acids in length. The structure of the gene, predicted from the nucleotide sequence, is shown schematically in Figure 3.7. The first nucleotide of putative translation initiation codon (ATG) is numbered +1. The sequence AAAAATGGC (-4 to +5) matches the plant consensus translation initiation sequence AACAATGGC (Joshi, 1987 and Lutcke, 1987). Two TATA-box like sequences were found at positions –187 to –184 for TATA and –144 to –137 for TATTTAAA. One putative CAAT box, CCCAAAT, was located at –272 to –266. The translation stop condon, TAA, was found at nucleotide position 2848 followed by a putative polyadenylation signal of TATAA at position 2913-2917 (Mclauchlan J, 1985). An intron of 711 bp was located at position of 390-1100. The predicted splice site of intron followed the conserved nucleotide /gu....ag/ rule for splice junctions (Brown, 1986). If this region was translated, there would be numerous stop codons present in all six reading frames.

```
-913  AAGCTTATTGTGTTATTTATTTTAATAATAATAATAATTTTTACCCAAAAAAACAACAAC   -854
-853  ATATTTAACTCTTACCAATACATTCTCACTTTATATTTATATTTATAACAAAATACTATA   -794
-793  AGAAATTTTAATAAAAATAAAAGATAAAACTATAGACTAAAATTAGCAAATTCGGTTAAT   -734
-733  TAGCTTCAATTTTTTAGTTATGTACAACACACAATTAATATTTTAATTACAAAAATATTT   -674
-673  AATTATATAATATAAAATTAAAATCTCATTGTTTAAATTTTTTATTATATGTTAAAAATT   -614
-613  ACATATTTTTATTTAATAAAAGGGTTTATATTTATATATTTTTTTATAAAATTATGTCAT   -554
-553  ATTTATACTAAATTAATTTATTAAAAATTAAAATACTTTTAATGAATATTTATTTTATTT   -494
-493  CTTCAAAAAAAAAGTTTCAATTTTACTTTAAAAAAAATCAAAATTTATTTTGAAATTTAC   -434
-433  TTACTGATTAGGTGGCTGTAGCGTTGGAGACCCAATCAATGGAAACGAATTTCTTGGCCA   -374
-373  GTTGCCCACCAACCACAACCTCACCATGCACCACCACCACGAGTCAAATTTACCCTTCTC   -314
-313  TCCTAATCATCACTCTCATGCAATCCCAACCCTTGGATTTTCCCAAATCAATGGCCATTA   -254
-253  TTAATTTTCAACCAACCCACCTTCTCTCCCTCTGCCCTCCTTATGCTTACCTACCCATTA   -194
-193  CACACTTATATGGAAGTCTCTTCACCTCTACTTCTCTCTCCCCACCTTCTATTTAAACTC   -134
-133  CACTCCTTCATCCTCTGCTCCTCAGGAAATCCATTTCCTACCAAAGATTTCCTCTCTCAG   -74
 -73  ATCCTTTGTTTCTCCTTCAACTTCAAGTTCTTCACTTCCTTGAGTTGTTTTGCTGTCTGG   -14
 -13  GTATTTCTTAAAAATGGCAACAATCTCTCAGAATGGTCACCAGAATGGTTCTTTAGACTC   47
                M  A  T  I  S  Q  N  G  H  Q  N  G  S  L  D  S
```


48 TTTGTGTACAGCTCGTGACCCACTGAATTGGGGTCTGGCTGCTGAGTCTATGAGTGGCAG 107
L C T A R D P L N W G L A A E S M S G S
108 CCATTTAGACGAAGTGAAAAAGATGGTGGCTGAGTTTAGGAAGCCTTTGGTCAAGTTAGG 167
H L D E V K K M V A E F R K P L V K L G
168 CGGCGAGACCTTGACTGTAGCTCAGGTGGCAGCTATTGCTCGTGAATCTGGTCTCCAAGT 227
G E T L T V A Q V A A I A R E S G L Q V
228 GGAGCTTGCAGAATCTGCTAGAGCTGGTGTTAAGGCGAGTAGTGACTGGGTCATGGATAG 287
E L A E S A R A G V K A S S D W V M D S
288 TATGAGCAAAGGAACTGATAGCTATGGTGTCACTACTGGATTTGGAGCCACTTCACATAG 347
M S K G T D S Y G V T T G F G A T S H R
348 AAGAACCAAGCAAGGTGGTGCTTTACAGAGGGAGCTCATTAGgtaaactttggttctctc 407
R T K Q G G A L Q R E L I R
408 ctccatattcaaatgtctaatagaaacaaaacgttaaaaaaaaatttcttcttctgctct 467
468 ttctacggtgggtggctcagtcgagttgactcagtcagttaaatggtgaatctatgggat 527
528 tgtcatttgagtaagcctggtgaaaattgagcaacctgactttatttttggtgaggccga 587
588 gttaactcagtcgagttaaatgagactcggctggtctttcatttgcatgagttgtcgtac 647
648 tcatgccggcctgacgagatatactatggcaaaaaaaaagtattttatggttacttattt 707
708 taaaaatttattatgttttatgaaaattatatttctaaaatggttttgaaaaacattttt 767
768 atgttataaatataaaaaagtaaaaagttttatctagcattaactaaatccaattttatt 827
828 acttaaatataaaaaaataaatatattaggatggatatggttggactcctgttcctgtag 887
888 ctgtaggtagatttgcccattttgtttgactacctgttaaactcacatttttattattaa 947
948 attagtaaaatttaattatttagttaatatattaatcaattgaaagagagatctaactaa 1007
1008 atgatacttttggcagccttagaattcatattttcctagaaaattcatccaaaacagtaa 1067
1068 ctaaacatgttgtatgttactttgtttctgcagATTCCTGAATGCTGGGATTTTTGGCAA 1127
F L N A G I F G N
1128 TAAGACAGAATCGTGTCACACTTTGTCACACTCTGCAACAAGAGCAGCAATGCTAGTGAG 1187
K T E S C H T L S H S A T R A A M L V R
1188 GATCAACACTCTTCTCCAAGGTTACTCAGGCATTAGATTTGAAATCTTGGAAGCTATCAC 1247
I N T L L Q G Y S G I R F E I L E A I T
1248 CAAGCTCCTCAACCACAATATTACTCCTTGCTTGCCGCTGAGAGGCACAATCACTGCTTC 1307
K L L N H N I T P C L P L R G T I T A S
1308 AGGGGATTTAGTTCCATTGTCCTACATTGCTGGATTGCTCACCGGCCGGCCTAATTCCAA 1367
G D L V P L S Y I A G L L T G R P N S K
1368 GGCTGTTGGCCCTAATGGAGAATCCCTAGATGCCCAGCAAGCCTTTCACTCTGCTGGTAT 1427
A V G P N G E S L D A Q Q A F H S A G I
1428 TGATTCTGGCTTCTTTGAGTTGCAGCCTAAAGAAGGCCTTGCTCTGGTTAATGGCACTGC 1487
D S G F F E L Q P K E G L A L V N G T A
1488 TGTTGGTTCTGGCTTGGCTTCCATGGTTCTCTTTGAGGCAAATGTTCTTGCTGTTTTATC 1547
V G S G L A S M V L F E A N V L A V L S
1548 AGAAGTCTTATCAGCTATTTTCGCCGAAGTTATGAATGGAAAACCTGAGTTTACTGATCA 1607
E V L S A I F A E V M N G K P E F T D H
1608 CTTGACTCATAAGTTGAAGCACCATCCAGGCCAAATAGAGGCTGCAGCTATAATGGAGCA 1667
L T H K L K H H P G Q I E A A A I M E H
1668 TATTTTAGATGGTAGCTCTTATATTAAAGCAGCTAAGAAGTTGCATGAAATTGATCCATT 1727
I L D G S S Y I K A A K K L H E I D P L

```
1728  GCAGAAACCAAAGCAAGATCGATATGCTCTCAGAACTTCCCCACAATGGCTAGGTCCTCA  1787
       Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G  P  Q
          .     .     .     .     .     .     .     .     .     .
1788  GATTGAAGTTATCAGATTCTCCACAAAATCGATCGAAAGAGAGATTAATTCAGTCAATGA  1847
       I  E  V  I  R  F  S  T  K  S  I  E  R  E  I  N  S  V  N  D
          .     .     .     .     .     .     .     .     .     .

1848  CAACCCTTTGATTGATGTTTCTAGGAACAAGGCCTTGCATGGTGGAAATTTCCAGGGGAC  1907
       N  P  L  I  D  V  S  R  N  K  A  L  H  G  G  N  F  Q  G  T
          .     .     .     .     .     .     .     .     .     .
1908  CCCAATTGGAGTCTCAATGGATAATGCACGTTTGGCCATTGCATCAATAGGAAAGCTCAT  1967
       P  I  G  V  S  M  D  N  A  R  L  A  I  A  S  I  G  K  L  M
          .     .     .     .     .     .     .     .     .     .

1968  GTTTGCTCAGTTCAGTGAGCTTGTAAATGATTTTTACAACAATGGGTTGCCATCAAATCT  2027
       F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S  N  L
          .     .     .     .     .     .     .     .     .     .
2028  CACAGCCAGCAGGAATCCAAGCTTGGATTATGGCTTCAAGGGAGCTGAAATTGCAATGGC  2087
       T  A  S  R  N  P  S  L  D  Y  G  F  K  G  A  E  I  A  M  A
          .     .     .     .     .     .     .     .     .     .
2088  TTCTTACTGTTCTGAGCTCCAATACCTTGCAAATCCAGTCACCAGCCATGTACAAAGTGC  2147
       S  Y  C  S  E  L  Q  Y  L  A  N  P  V  T  S  H  V  Q  S  A
          .     .     .     .     .     .     .     .     .     .
2148  AGAGCAGCACAATCAAGATGTAAATTCATTGGGGCTAATTTCTTCAAGAAAGACAGAAGA  2207
       E  Q  H  N  Q  D  V  N  S  L  G  L  I  S  S  R  K  T  E  E
          .     .     .     .     .     .     .     .     .     .
2208  AGCTGTAGACATCTTGAAGCTCATGTCCACGACTTTCTTAGTAGCACTTTGTCAAGCTAT  2267
       A  V  D  I  L  K  L  M  S  T  T  F  L  V  A  L  C  Q  A  I
          .     .     .     .     .     .     .     .     .     .
2268  TGACTTGAGGCATTTGGAGGAGAACTTGAAGCACGCAGTCAAAAACACAGTAAGCCAAGT  2327
       D  L  R  H  L  E  E  N  L  K  H  A  V  K  N  T  V  S  Q  V
          .     .     .     .     .     .     .     .     .     .
2328  AGCTAAGAGGATTCTAACTACAGGAGCTAGTGGAGAACTTCACCCATCAAGATTCTGCGA  2387
       A  K  R  I  L  T  T  G  A  S  G  E  L  H  P  S  R  F  C  E
          .     .     .     .     .     .     .     .     .     .
2388  GAAGGACTTGCTCAAAGTGGTGGATCGCGAGCAAGTCTTCTCTTATGTCGATGACGCCTG  2447
       K  D  L  L  K  V  V  D  R  E  Q  V  F  S  Y  V  D  D  A  C
          .     .     .     .     .     .     .     .     .     .
2448  CAGTGCTACCTATCCATTGATGCAAAAACTAAGGCAAGTTCTCGTGGACCATGCCTTGGC  2507
       S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  H  A  L  A
          .     .     .     .     .     .     .     .     .     .
2508  AAATGGCGAGAGTGAGAAGAATGCCAGCACTTCAATCTTCCAAAAGATCAGAGCTTTTGA  2567
       N  G  E  S  E  K  N  A  S  T  S  I  F  Q  K  I  R  A  F  E
          .     .     .     .     .     .     .     .     .     .
2568  GGAAGAATTGAAAGCCCTTTTGCCTAAAGAAGTTGAGAGTGCAAGAGAGGCATACGAGAA  2627
       E  E  L  K  A  L  L  P  K  E  V  E  S  A  R  E  A  Y  E  N
          .     .     .     .     .     .     .     .     .     .
2628  CGGGAATCCAGCAATTGCCAACAAGATCAAGGAATGCAGATCTTACCCATTGTATAAGTT  2687
       G  N  P  A  I  A  N  K  I  K  E  C  R  S  Y  P  L  Y  K  F
          .     .     .     .     .     .     .     .     .     .
2688  TGTGAGAGAGGAAATAGGAACTGGGTTGCTCACCGGAGAAAAGATCCGGTCACCGGGAGA  2747
       V  R  E  E  I  G  T  G  L  L  T  G  E  K  I  R  S  P  G  E

          .     .     .     .     .     .     .     .     .     .
2748  GGAATTTGATAAGGTTTTCACTGCTATGTGCCAAGGAAAGATCATTGATCCAATGCTGGA  2807
       E  F  D  K  V  F  T  A  M  C  Q  G  K  I  I  D  P  M  L  D
          .     .     .     .     .     .     .     .     .     .
2808  TTGTCTCAAAGAGTGGAATGGTGCCCCTCTTCCAATATGTTAAACTGTAACTTTCTTGTT  2867
       C  L  K  E  W  N  G  A  P  L  P  I  C  *
          .     .     .     .     .     .     .     .     .     .
2868  TTGTTTACACTTCAAGATTTGTTTTCCAATTGCTTTTTATGTACTTATAATTTGTGATGT  2927
          .     .     .     .     .     .     .     .     .     .
2928  AAAAAATCTGTAATGCATTTCTTTTAAATGTTCAATTGTTATCTTCTCACTTTTGTGCTG  2987
          .     .     .     .     .     .     .     .     .     .
2988  GAATTGAAGGCAGAATAGCAATGGTAATTACTTCAAGATTACTGAATTGAAATTTTTTGA  3047
          .     .     .     .     .     .     .     .     .     .
3048  GTTATTCATTAATCTAAGTGTTTATTTAAAAAAATTTAAGAGAATTTTAATTGACAGAAA  3107
          .     .     .     .     .     .     .     .     .     .
3108  GAAGTGTATCCATATATTTTCGAGGTCGAGAATCCATGAACGTGTAGCGGTTGGCGCTGT  3167
          .     .     .     .     .     .     .     .     .     .
3168  GCTGTTCCTCTCCTTGACCTTGAGCTTATTGCTTGGCAGCTTGAGAGCCTCTCCTTGAAG  3227
          .     .     .     .     .     .     .     .     .     .
3228  GGAAGACGAGAGATGTACCTCTTCTTGTTTGGATATATGGGGCCCATTAGATTTCCCAAA  3287
          .     .     .     .     .     .     .     .     .     .
3288  ATTAATGTTTCGTTGTTTCTACGTACTTCTGTTTTCACCCCAGATGTTGATGCTTAACAA  3347
```

```
3348  ACTCCAAATTTCAATTGAAATTAATTTTTTTTATTGTGCATATATGGTAAAAAAAAAGAA  3407
3408  TATAAAAAAATAAAATTAGAATATAATTAAATAAGAAGTTTTTTTTGAAATTGAAAATTT  3467
3468  TACTATATCTATCATTAATAATGATTTGGTCTTTGTGTTTTAAAACATTAGACTGTTTAG  3527
3528  TTCCTATTTTCTTTTGCCATACACTTGGATCGTTTTATCCATTTTTCCTTTAGTATGGTA  3587
3588  ATATAAAATTATATTTAGTCCATAGATATTTATTTAAAATTCCTCATAATTATTTACAAT  3647
3648  TTCATATATTAATTCTTCTTAGAGCATAGAGATATATAAAAAGAGGGAGAGAAAATTAAG  3707
```

Figure 3.6 Complete sequence of the MePAL2 gene and its deduced amino acid sequence. Translation initiation and termination sites are at position 1 and 2846 in bold. The intron region (size 711 bp) is in lower case. The exon/intron boundaries are in bold. Two putative TATA boxes and CAAT box are in bold and underlined.

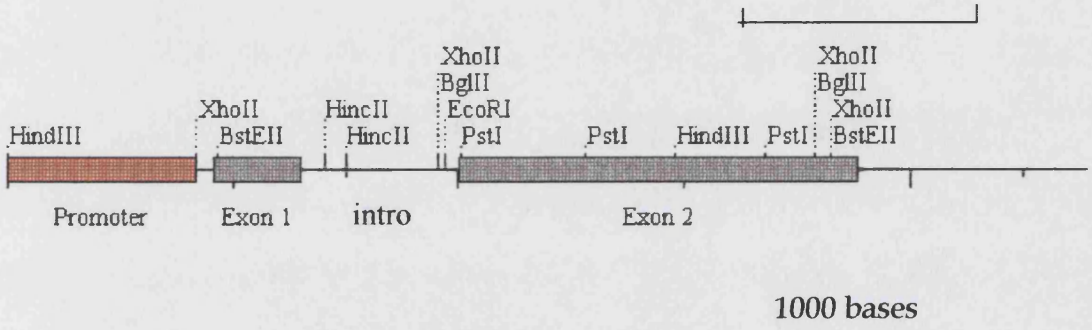


Figure 3.7 The structure of the MePAL2 gene.

The intron and 3'untranslated regions are represented by straight lines. The promoter region is indicated by a red box whilst two putative exons by grey boxes.

3.6 Extending the promoter sequence of MePAL2 by sub-cloning a *Bgl*II fragment

Most *cis*-elements of promoters are within the 500 bp upstream of the transcription start site; however, some of them are beyond 1000 bp. Initially, only 913 bp of the 5' UTR sequence of MePAL promoter was sequenced. Therefore, to gain more upstream sequence, the restriction map of the known sequence of MePAL2 and the *Bgl*II restriction map of clone 1A were analysed. In the MePAL2 restriction map, there are two *Bgl*II sites, at 995-1001 and 2665-2670 (Figure 3.6 and Figure 3.7). The *Bgl*II digestion of clone (Figure 3.4) 1A phage DNA released two fragments that hybridized to the 1.9 kb *Eco*RI probe from cDNA MePAL1. One fragment was 1.6 kb, which must corresponded to the *Bgl*II fragment generated by the 995 bp to 2670 bp site. Another fragment was about 4-5 kb, which might be a fragment released by the 995 bp-1001 bp *Bgl*II site to a possible upstream *Bgl*II site or by the 2265 bp to 2270 bp *Bgl*II site and a

possible downstream *Bgl*II site. If the later assumption were true, the 4 kb fragment would not bind to the 1.9 kb *EcoR*I MePAL1 probe since the probe did not cover this region. Therefore, the 4 kb fragment must be upstream from 995 bp –1001 bp of *Bgl*II site, and contain more sequence of the promoter. To extend the MePAL2 5'uptream region, the phage DNA of clone 1A was digested with *Bgl*II, and the resulting 4 kb fragment was purified, blunt ended and subcloned into the *Sma*I site of the pUC19 plasmid. The insert was sequenced by universal primers and specific primers which were designed from identified sequence. Figure 3.8 shows the extended promoter sequence of MePAL2, which is 2740 bp in length from the ATG translation start site.

```
-2740  CTTCGTTTCCCACAGACGGCGCCAATTGATGATCTGAGATCCAATAGATTAGGGTTTTAC  -2681
-2680  AAGGGTTTTGTAATGAATAATGAAAGCTTAAACTTCCTGGGGAGGGGACTCCCCTTTTAT  -2621
-2620  ACATTGTCTTGCTTGCTGGTGACGTGTAAAGGCCTCGCCATGATTGGGACACGCGTCCCT  -2561
-2560  GACATACAGATTCGGGCGTACGAGGGTATCAGCCGCCTGCTCCATGCGTAACGGCCTCTG  -2501
-2500  ATTCCCCCCTGGGCGTACGTTTGTGCGGACCTGCTAGGCTGTCTGGTGAGTACCATTCTG  -2441
-2440  GGTTCTGGGCCGGACTGGAGAGTGGGCCTGATGTTGGGCCTGGGTGGAGAGGATCTGCTC  -2381
-2380  GAGGGAGTATTCGGGCTGGACCATCCTGGGTGAGAAGTCGTAGCCGAGCCCGGCCTCCGG  -2321
-2320  CGCTGAGGGAGCTGGACCTGCTTCATGGCCAGGTGCATAGGCCTTCTCTCCATGGGCTTT  -2261
-2260  TCGTTGTGGGCTAGGCCCACGATGGGGATAAGGAAATCCAGCGGTCATCAAAAATTAATA  -2201
-2200  AGAGAAATATAAAGAAAGGATAAGATTGGTAATTAAGCCAAATGAATACAAAGGTTTCAT  -2141
-2140  GACAGAGAATAAAGGTTTCTAACTCCTTTTTAAAAAACAAAATTCATTATGTAATCTTCT  -2081
-2080  CTCTTTTGTGTAAGGTTTCATAGTAAAAATAAAAATTTATTTCATAAAATTCATTATATA  -2021
-2020  AATAATTAAGTAAAATTATTAATCTATAAAAATTTATAATACCTAAGTAAAATTAAGAAT  -1961
-1960  ATAATTACTGATAACCGTCGGATATTCATTTTATTGGATCAGGGCGAATCCAAAAAAGTA  -1901
-1900  GAATCAGTTCAAAACTTAGTAGATACTCTAGAATCACCCAATCCGACCCTTTTAAGTTGG  -1841
-1840  GCTCGAGGGCATGAAAAGGGAGTCAGATCGCATATGTGCTTTAACTCAGCATGTTAAAGA  -1781
-1780  TCGACTCAGTGACGTGAAAAGAAGTGAGATAGCAGATCCGGTCACTTTATAACCGTTTTC  -1721
-1720  GTTTATACGCCAAAAAATATTAAATGACCGCCTGATACGGAATGAATCTCTGGATATCCG  -1661
-1660  TACACGTGGGTCAGTGTGTCAGGACAAATGGTATTATAACAGAAAGACCATTACACATTT  -1601
-1600  AAATAAAAAGAAAAAAAAATAAGACGAAAAAAGAATTATCTAGAAGGACCAAATTTACAC  -1541
-1540  ATACATACTAAAAATCTTTTTATTCTCTTAGTCAGAACATCAATTACTTTTTATAAATTT  -1481
-1480  AGAGATTAAATAGTTTATTATAATTTAATCAGTAATTTTTTTTTATATATTATTTTAAAG  -1421
-1420  CAATATTTGGTTAGTGTTTATAATATATTGACCATTATTTTGGCTTCTTATTACAAGTAG  -1361
-1360  AGCCATTGCTAGCTAATGACAATATTCTATTATTAATGGTCAACAAAGGTAAGCTTAAAG  -1301
-1300  CTAAAGTCCTTAATCAATTGCTCAGATTTTATCAAAATAACTTCAATTCTTAAGAGATAC  -1241
-1240  TATTCCCACCTACCATCACTCTATTTTTCTTTCTTTTCTTTTTTCTTTTTTAGATTTTTA  -1181
```

```
-1180  ATCACTGCCACCTAAACTATGTAATCTCTCATGGAAAATATTTTTTAACACATAAAAAAT   -1121
-1120  AAAAAAATAGAAAATTACACATTCATTATAATCTAAATATCAGTAAAACATAGAAAATTT   -1061
-1060  AATTACTGAAATTTTATTTAACATAGAATAAAATTTAATTAATTTAGTGATGTTTCTCAC   -1001
-1000  TTCTCAGCCTCCAATGGCAAAGATCGATCCAAAGAAATCATAGCTGTCCATGTTAGGGTT   -941
 -940  CCTCAAATCTCGTGTATATTTTAAACTAAGCTTATTGTGTTATTTATTTTAATAATAATA   -881
 -880  ATAATTTTTACCCAAAAAAACAACAACATATTTAACTCTTACCAATACATTCTCACTTTA   -821
 -820  TATTTATATTTATAACAAAATACTATAAGAAATTTTAATAAAAATAAAAGATAAAACTAT   -761
 -760  AGACTAAAATTAGCAAATTCGGTTAATTAGCTTCAATTTTTTAGTTATGTACAACACACA   -701
 -700  ATTAATATTTTAATTACAAAAATATTTAATTATATAATATAAAATTAAAATCTCATTGTT   -641
 -640  TAAATTTTTTATTATATGTTAAAAATTACATATTTTTATTTAATAAAAGGGTTTATATTT   -581
 -580  ATATATTTTTTTATAAAATTATGTCATATTTATACTAAATTAATTTATTAAAAATTAAAA   -521
 -520  TACTTTTAATGAATATTTATTTTATTTCTTCAAAAAAAAAGTTTCAATTTTACTTTAAAA   -461
 -460  AAAATCAAAATTTATTTTGAAATTTACTTACTGATTAGGTGGCTGTAGCGTTGGAGACCC   -401
 -400  AATCAATGGAAACGAATTTCTTGGCCAGTTGCCCACCAACCACAACCTCACCATGCACCA   -341
 -340  CCACCACGAGTCAAATTTACCCTTCTCTCCTAATCATCACTCTCATGCAATCCCAACCCT   -281
 -280  TGGATTTTCCCAAATCAATGGCCATTATTAATTTTCAACCAACCCACCTTCTCTCCCTCT   -221
 -220  GCCCTCCTTATGCTTACCTACCCATTACACACTTATATGGAAGTCTCTTCACCTCTACTT   -161
 -160  CTCTCTCCCCACCTTCTATTTAAACTCCACTCCTTCATCCTCTGCTCCTCAGGAAATCCA   -101
 -100  TTTCCTACCAAAGATTTCCTCTCTCAGATCCTTTGTTTCTCCTTCAACTTCAAGTTCTTC   -41
  -40  ACTTCCTTGAGTTGTTTTGCTGTCTGGGTATTTCTTAAAA
```
ATG

Figure 3.8 The extended 2740 bp MePAL2 promoter sequence. The translation intiation site ATG is in bold.

3.7 Isolation of genomic MePAL1

In addition to p1A, the other three positive MePAL clones p2A, 3E and 4A, were confirmed by PCR and sequenced in certain regions. Unfortunately all of them were from the same gene as genomic MePAL2 (3E showed exactly the same sequence as MePAL2 but it is shorter than the other clones) and none of them corresponded to MePAL1 or MePAL3 cDNA. In order to isolate a MePAL1 genomic clone, a gene specific probe was needed to screen cassava genomic library. Another approach would be to amplify genomic DNA by PCR using primers specific to MePAL1 cDNA.

3.7.1 Screening cassava the genomic library with a specific probe

PAL genes are highly similar throughout their coding regions (see Section 3.8). The exon sequences in particular show high similarity to one another, ranging from 70 to 90 per cent identity (Han, 2000), especially in the second exon (Wanner, 1995). However, the 5' UTR and the 5' end of coding region of PAL genes often show divergence (Figure 3.14). It has been shown that most PAL introns are not only different in size but also have quite divergent sequences (Section 3.8.4). Therefore, these divergent regions could be used as gene-specific probes. In order to retrieve gMePAL1 from the genomic library, MePAL1-specific probes made from the 5' region and intron were used to screen the genomic library.

Firstly, a 180 bp *Bam*HI / *Pst*I fragment (Fig. 3.1A) from the 5' end of MePAL1 cDNA was used as a specific probe, since there was no similarity between MePAL2 and MePAL1 within this 180 bp region except that there was 78% similarity for the last 34 bp of the fragment. The specificity of the probe was tested for its binding to cDNA MePAL1 and MePAL2 in hybridisation. As it can be seen (Figure 3.9) the probe bound to MePAL1 cDNA but not to MePAL2 genomic DNA. About $2.3x10^8$ plaques of the genomic library were screened with the specific probe but no positive clone was isolated. The screening experiment was repeated twice and no positive was identified.

In almost all the introns of plant PAL genes have been studied, the positions were conserved (Wanner, 1995) but the sequence of introns were very divergent from one and another (Section 3.8.5 discussion). When the sequence of genomic MePAL1 was determined using PCR amplification later on, the intron (598 bp) of MePAL1 was quite divergent from that (711 bp) of MePAL2 in most of the region. Therefore, the intron from MePAL1 (see Section 3.7.2), amplified by PCR (primers used GPA1F and GPAL1R11 see Section 3.7.2), was used as a specific probe to isolate the genomic clone corresponding to MePAL1 cDNA from genomic library. About 2.3×10^8 plaques of the genomic library were screened three times with the intron from MePAL1 as a specific probe, but no positive clone was isolated.

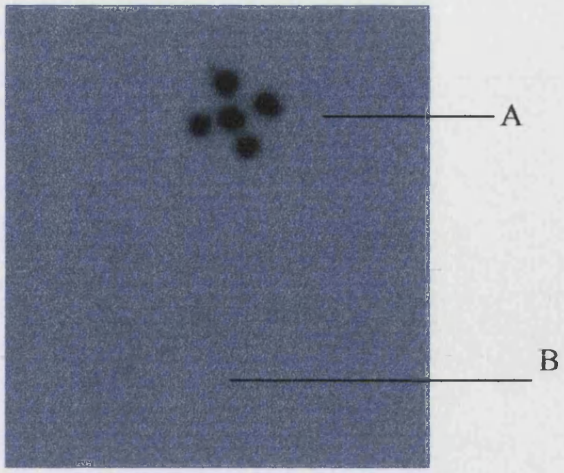


Figure 3.9 Specificity of the 180 bp ***Bam*HI / *Pst*I MePAL1 specific probe.** A is cDNA MePAL1 clone while B is genomic MePAL2 clone. Hybridization stringency was 60°C overnight and final wash is 0.1 x SSC / 0.1% of SDS at 60°C.

The failure to isolate MePAL1 from the genomic library may have been due to there being no such clone in the library screened. The clone might have been lost during the amplification from the original library as the library screened was an amplified genomic one.

3.7.2 Amplifying genomic MePAL1 by PCR

Three sets of primers, GPA1F/GPA1R, GPAL1F2/GPAL1FR2 and GPA1F/GPAL1R11, were designed from different regions of MePAL1 cDNA, which had low similarity to genomic MePAL2 sequence, in order to amplify the relevant region from genomic DNA (Figure 3.10).

These three sets of PCR primers are:

1. GPA1F: TTTCTCACCACTCCTCTACC (located at 34-53 of MePAL1 cDNA)
   GPA1R: GAAGCAATG GCTAAACGAGTG (located at 1312-1292 (r) of MePAL1 cDNA)
2. GAPL1F2: CCAATGTGTTGGCAGTCCTG (located at 886-905 of MePAL1 cDNA)
   GPAL1R2: GGATCAATCAGCTTCCCTGC (located at 2158-2139 (r) of MePAL1 cDNA)
3. GPA1F: TTTCTCACCACTCCTCTACC (located at 34-53 of MePAL1 cDNA)
   GPAL1R11: TAAGCCTGAACCAACTGCAG (located at 863-844 (r) of MePAL1 cDNA)

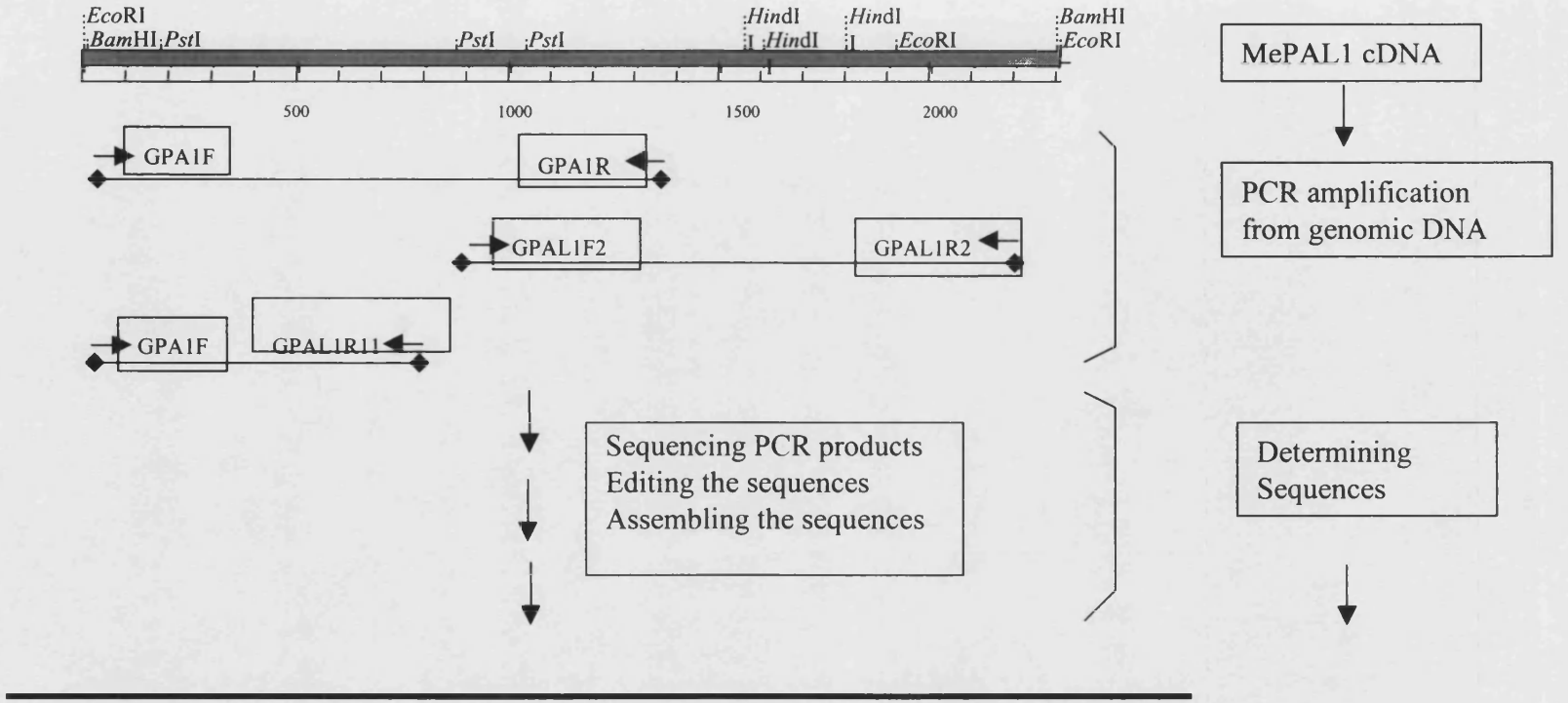


Figure 3.10 Schematic map of amplification and sequencing of MePAL1 genomic DNA.

Three sets of primers, GPA1F/GPA1R, GPAL1F2/GPAL1FR2 and GPA1F/GPAL1R11, were designed from different regions of MePAL1 cDNA, to amplify the relevant region from genomic DNA by PCR. Three PCR products were then sequenced respectively by the primers used for PCR or primers designed from both directions based on the sequenced regions. Finally the sequences from those three PCR products were assembled by GCG program (see section 2.5 in Chapter 2).

Cassava genomic DNA from cultivar MNGA1, from which MePAL1 cDNA was isolated, was used as a template for the PCR amplification. The PCR was carried out with one cycle of 94°C 3 minutes, followed by 35 cycles of 94°C 1 minute, 55°C 1 minute, 72°C 3 minutes, and completed with 72°C for 10 minutes. The PCR products were purified and sequenced. The assembled MePAL1 genomic sequence had 100% identity to MePAL1 cDNA except for a 598 bp genomic DNA region, which was identified to be an intron region. The position of the intron in the gene was similar to that of the gMePAL2 intron and the size of the intron was smaller than the MePAL2 intron (Figure 3.11). The exon-intron junction followed the sequence **/gu……ag/…**conserved in plant systems (Brown, 1986).

```
                                        GPA1F  →
   1 CTCAAACTTCCTATACCTATATTTTTCTCCCACTTTCTCACCACTCCTCTACCTTTCCGG   60
  61 CAAAATAAATCAGCAAAATGGAGTTTTGTGAGGCTCACAATGTAACTGCTTCACCTGGAT  120
                     M  E  F  C  E  A  H  N  V  T  A  S  P  G  F
 121 TCTCCTCTGCTGACCCATTGAACTGGGGCATGGCTGCAGAGTCACTGAAGGGCAGCCACC  180
       S  S  A  D  P  L  N  W  G  M  A  A  E  S  L  K  G  S  H  L
 181 TTGATGAGGTGAAGCGCATGGTGGATGAGTACAGGAAGCCTGTGGTGAGGCTAGGGGGTG  240
       D  E  V  K  R  M  V  D  E  Y  R  K  P  V  V  R  L  G  G  E
 241 AGACCCTGACTATAGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGTTG  300
       T  L  T  I  A  Q  V  T  A  I  A  N  H  D  S  G  V  K  V  E
 301 AGCTGTCTGAGGAGGCTCGAGCTGGGGTCAAGGCCAGTAGTGACTGGGTTCTTGATTCCA  360
       L  S  E  E  A  R  A  G  V  K  A  S  S  D  W  V  L  D  S  M
 361 TGAATAAAGGAACCGATAGTTATGGTGTCACCACTGGCTTTGGTGCAACCTCCCATAGAA  420
       N  K  G  T  D  S  Y  G  V  T  T  G  F  G  A  T  S  H  R  R
 421 GAACCAAGCAGGGGGGTGCCCTTCAGAGAGAACTCATA
       T  K  Q  G  G  A  L  Q  R  E  L  I

                                                 aggtaacaaatattaaaaatat  480
 481 ataaatatttataaactagttgattattttgagatcatccctataatagactagatccat  540
 541 aatggataaaatgatgaaagtacaagtttctttttctttttcttaaaagggtaatttaaaa  600
 601 aaatatttttatcagaaaataaataaataaagagtacacgaaacgaaactgccttctact  660
 661 aagtctaaatgcaaatcatctaccccaacagtactgttaattaataatctaaattatttt  720
 721 acagtattttcaccatcccattaaaatagatatattcatttttttatataatttaaaaaa  780
 781 tataattaaaaataatacatttatattgccttttttagtaatatatacattactcatatta  840
 841 ctcaattattaaatcatttttttaaattgataaattttattttttattatatattaaataa  900
 901 gggtatattaaaaatattaaatgatagattcatataaatagctaattccaacaaatctga  960
 961 aattaaaaatttaattgacttgagcattgtttcttgaacaaatggttagtaatttaaatg 1020
1021 tttttcttggaaatttttgttttattctcc

                                        AGATTCTTGAATGCTGGGATCTTTGGAAAT 1080
                                        R  F  L  N  A  G  I  F  G  N
1081 GGACAAGAATCTTGCCACACATTGTCTCACACTGCAACTAGAGCAGCAATGCTGGTGAGA 1140
     G  Q  E  S  C  H  T  L  S  H  T  A  T  R  A  A  M  L  V  R
```

```
1141  ATCAACACCCTCCTGCAAGGTTATTCAGGCATAAGATTTGAAATCCTGGAAGCCATTACC  1200
      I  N  T  L  L  Q  G  Y  S  G  I  R  F  E  I  L  E  A  I  T
1201  AAGTTCATCAACAACAATGTTACTCCGCGTTTGCCCCTCAGAGGCACAATCACAGCCTCT  1260
      K  F  I  N  N  N  V  T  P  R  L  P  L  R  G  T  I  T  A  S
1261  GGTGACCTGGTCCCACTGTCCTACATTGCCGGGCTTTTGACCGGCCGGCCCAACTCCAAG  1320
      G  D  L  V  P  L  S  Y  I  A  G  L  L  T  G  R  P  N  S  K
1321  TCGTTAGGGCCCAATGGAGAATCCTTGGATGCAGCTGAAGCCTTTAAGCTTGCTGGGATC  1380
      S  L  G  P  N  G  E  S  L  D  A  A  E  A  F  K  L  A  G  I
1381  AATGGTGGATTTTTTGAATTGCAGCCCAAAGAGGGTCTAGCTTTAGTAAATGGTACTGCA  1440
      N  G  G  F  F  E  L  Q  P  K  E  G  L  A  L  V  N  G  T  A

      ← GPAL1R11                                  GAPL1F2 →
1441  GTTGGTTCAGGCTTAGCTTCTATGGTTCTTTTTGAGGCCAATGTGTTGGCAGTCCTGTCA  1500
      V  G  S  G  L  A  S  M  V  L  F  E  A  N  V  L  A  V  L  S
1501  GAAGTTTTATCAGCAATTTTTGCAGAAGTTATGCTTGGAAAACCAGAGTTTACAGATCAT  1560
      E  V  L  S  A  I  F  A  E  V  M  L  G  K  P  E  F  T  D  H
1561  TTGACACATAAATTGAAGCATCATCCTGGACAAATTGAAGCTGCAGCAATTATGGAACAT  1620
      L  T  H  K  L  K  H  H  P  G  Q  I  E  A  A  A  I  M  E  H
1621  GTCTTGGATGGAAGTTCTTATATTAAAGCAGCTCAAAAGGTTCATGAAATTGATCCATTG  1680
      V  L  D  G  S  S  Y  I  K  A  A  Q  K  V  H  E  I  D  P  L
1681  CAGAAGCCTAAACAAGACAGATATGCTCTTAGAACATCTCCTCAATGGCTTGGCCCTCAG  1740
      Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G  P  Q
1741  ATTGAAGTGATTCGAACAGCGACTAAAATGATCGAACGCGAAATCAACTCTGTGAATGAT  1800
      I  E  V  I  R  T  A  T  K  M  I  E  R  E  I  N  S  V  N  D
1801  AATCCATTGATTGATGTCTCCAGGAATATTGCTTTACATGGAGGCAATTTCCAGGGGACC  1860
      N  P  L  I  D  V  S  R  N  I  A  L  H  G  G  N  F  Q  G  T

                                        ← GPA1R
1861  CCAATTGGTGTTTCAATGGATAACACTCGTTTAGCCATTGCTTCAATTGGTAAACTCATG  1920
      P  I  G  V  S  M  D  N  T  R  L  A  I  A  S  I  G  K  L  M
1921  TTTGCTCAATTCTCTGAGCTTGTTAATGATTTTTACAACAATGGGTTGCCTTCAAATCTC  1980
      F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S  N  L
1981  ACTGGTGGACGCAATCCAAGCTTGGATTATGGCTTCAAAGGAGCTGAAATTGCCATGGCA  2040
      T  G  G  R  N  P  S  L  D  Y  G  F  K  G  A  E  I  A  M  A
2041  TCTTACTGCTCAGAGCTCCAATTTCTTGCCAATCCTGTAACTAATCATGTCCAAAGTGCA  2100
      S  Y  C  S  E  L  Q  F  L  A  N  P  V  T  N  H  V  Q  S  A
2101  GAGCAGCACAACCAAGATGTTAACTCACTAGGCTTGATTTCTTCAAGGAAAACAGCTGAA  2160
      E  Q  H  N  Q  D  V  N  S  L  G  L  I  S  S  R  K  T  A  E
2161  GCTGTTGACATATTGAAGCTCATGTCTTCTACATACTTAGTTGCTCTATGTCAAGCCATT  2220
      A  V  D  I  L  K  L  M  S  S  T  Y  L  V  A  L  C  Q  A  I
2221  GACTTGAGACACTTGGAGGAGAACTTGAAGCAAACAGTCAAGAACACAGTAAGTCAAGTT  2280
      D  L  R  H  L  E  E  N  L  K  Q  T  V  K  N  T  V  S  Q  V
2281  GCAAAGAGAGTCTTGACAATGGGCATCAACGGCGAGCTCCACCCGTCGAGATTCTGCGAA  2340
      A  K  R  V  L  T  M  G  I  N  G  E  L  H  P  S  R  F  C  E
2341  AAAGACCTTCTCAAAGTCGTCGACAGGGAATACGTTTATGCATATGTTGATGATCCTTGC  2400
      K  D  L  L  K  V  V  D  R  E  Y  V  Y  A  Y  V  D  D  P  C
2401  AGTGCAACATACCCATTAATGCAAAAGCTGAGACAAGTACTAGTTGATCATGCCATGATG  2460
      S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  H  A  M  M
2461  AATGGTGAAAAGGAGAAGAATTCAAGCACTTCCATTTTCCAAAAAATTGGAGCCTTTGAA  2520
      N  G  E  K  E  K  N  S  S  T  S  I  F  Q  K  I  G  A  F  E
```

```
2521  GAAGAACTCAAGACCCTTTTGCCTAAAGAAGTAGAAAGTGCAAGAACTGAATATGAGAAT  2580
      E  E  L  K  T  L  L  P  K  E  V  E  S  A  R  T  E  Y  E  N
2581  GGTAATCCAGCTATTTCTAACAAGATCAAAGAATGTAGGTCATATCCACTATACAAGTTT  2640
      G  N  P  A  I  S  N  K  I  K  E  C  R  S  Y  P  L  Y  K  F
2641  GTGAGGGAAGAACTCGGTTGTAGTTTACTGACCGGCGAGAAGATTCGATCGCCCGGCGAA  2700
      V  R  E  E  L  G  C  S  L  L  T  G  E  K  I  R  S  P  G  E

                                          ◄— GPAL1R2
2701  GAGTTTGATAAGGTTTTCTCAGCAATTTGTGCAGGGAAGCTGATTGATCCCATGCTTGAA  2760
      E  F  D  K  V  F  S  A  I  C  A  G  K  L  I  D  P  M  L  E
2761  TGCCTCAAGGAGTGGAATGGTGCTCCTCTTCCAATCTGCTAAGTTTATTTTTTTTTGTTT  2820
      C  L  K  E  W  N  G  A  P  L  P  I  C  *
2821  GTTTGCTCATAAAGTCTACACATGC  2845
```

Figure 3.11 Nucleotide sequence of genomic MePAL1 and its deduced amino acid sequence.
Translation initiation site (ATG) and termination sites (TAA) are in bold. The intron region is in lower case whilst the exon and untranslated 5' and 3' are in upper case. The exon/intron boundaries are in bold. Peptide starts at 78; Intron from 459 to 1050; Coding region ends at 2802.
The PCR primers for isolating MePAL1 genomic DNA are underlined.

3.8 Discussion

3.8.1 Cassava PAL gene organisation

PAL genes have been isolated from a number of plant species and in most species studied PAL is encoded by a multigene family. It was estimated by Southern analysis or proved by the isolation of PAL genes that there are at least four PAL genes in rice (Zhu, 1995), three PAL genes in *Arabidopsis* (Wanner, 1995), less than 6 in alfalfa (Gowri, 1991), two in pea (Yamada, 1992), three in French bean (Shufflebottom,1993), four in parsley, at least two in poplar (Subramaniam,1993), four in tobacco (Fukasawa-Akada,1996), and five different classes in tomato (Lee, 1992). The complexity of the PAL gene family was demonstrated by studies of PAL genes in potato and pine tree. In potato, it was estimated that haploid genome contained 40 to 50 PAL genes (Joos *et al.*, 1992). At least ten cDNA clones representing ten different PAL genes were isolated from a cDNA library of potato suspension cells treated with fungal elicitor, indicating that these genes were active (Joos *et al.*, 1992). Restriction maps of 12 selected PAL genomic clones also showed considerable heterogeneity (Joos *et al.*, 1992). Based on

the tetraploid nature of potato, it might be argued that various genomic or cDNA clones might represent allelic forms, however, only three out of eight possible alleles were detected for the two 4-coumarate: CoA ligase genes from the same potato cultivar as used for the analysis of PAL genes (Joos *et al.*, 1992). In pine trees, *Pinus taeda* was distinguished by its PAL being encoded by a single gene (Whetten and Sederoff, 1992), whereas a diverse family of PAL genes were expressed in jack pine, *Pinus banksiana,* and its cell cultures (Butland *et al.*, 1998). Using degenerate PCR primers targeting the conserved sequences in PAL genes, Butland *et al* (1998) showed that PAL in jack pine was encoded by at least 8 to 10 loci. Five classes of PAL sequences were distinguished among 28 clones obtained from PCR amplification of haploid genomic DNA, which shared 68.8% to 94% nucleotide identity over 366 bp region compared. Using probes from jack pine PAL genes, it was also revealed that multiple PAL genes were expressed in *P. contorta.* Butland *et al* (1998) further suggested that *P. taeda* PAL must also be encoded by a similar multigene family. In fact, Whetten and Sederoff (1992) noted weaker bands in Southern analysis of genomic DNA hybridised with a PAL probe, suggesting the possible presence of diverged copies of the gene.

Southern hybridisation of cassava genomic DNA with a 1.9 kb MePAL1 fragment as a probe suggested that PAL was also encoded by a multigene family; multiple bands were observed when genomic DNA was digested with different restriction enzymes including *Eco*RI, *Bgl*II, *Xba*I and *Hin*dIII (Figure 3.1). Some of the bands were strong and some were weak, which indicates their different similarity to the PAL probe used and may represent the diverged PAL gene family members in cassava. *Eco*RI and *Bgl*II digested DNA showed two bands over 8 kb with strong signals, which may represent the MePAL1 gene and another PAL gene with high similarity to the 1.9 kb MePAL1 probe. In these two digestions, at least two more bands were observed. These results suggest that there may be at least four PAL genes in cassava, which was also supported by the Southern pattern of *Hin*dIII digestion. MePAL1 can be cut by *Hin*dIII, and when genomic DNA was cut with *Hin*dIII, the number (8 bands ranged from 0.7 to 6.5 kb) of the bands observed by hybridisation doubled that of *Bgl*II. *Xba*I digestion did not release so many bands binding to the probe, but the sizes of the bands were quite large and the signal was strong, which may indicate that PAL genes were clustered or *Xba*I released fragments (containing PAL sequence) of similar size.

3.8.2 Isolation and characterization of two cassava PAL genes

In this chapter I describe the organization and isolation of cassava phenylalanine ammonia-lyase genes. Genomic Southern blotting indicated that there were at least four PAL genes in cassava and two of them, genomic MePAL1 and MePAL2, were isolated. MePAL2 was isolated from a genomic library using a MePAL1 cDNA probe and sequenced. The whole sequenced region was 6447 bp which including 2741 bp of 5'upstream region, two exons of 388 bp and 1748 bp interrupted by one 711 bp intron, and 860 bp of 3' untranslated region. It encoded a polypeptide of 712 amino acids. The MePAL1 gene was isolated by PCR amplification using genomic DNA as a template but the 5' flanking region was not isolated though effort was made to isolate a MePAL1 genomic clone from the genomic library using a MePAL1 specific probe. The sequence region was 2846 bp in length and consisted of two 381 bp and 1748 bp exons and one 592 bp intron. The deduced amino acid sequence was 710 AA in length. As expected, in the coding region two cassava PAL genes were highly similar, especially in the second exon, in which the nucleotide similarity was 89%.

3.8.3 Similarity among cassava PAL genes

3.8.3.1 Similarity of DNA sequence in the coding region

The coding region sequence of MePAL2 was compared to MePAL1, MePAL3 (partial sequence) and MePAL (partial sequence). The similarity between MePAL1 and MePAL2 was 79%. MePAL2 and MePAL3 were very close in nucleotide sequences, about 98% similarity, with only 15 nucleotide difference over 1000 bp (Han, 2000). The similarity was 83% between MePAL2 and MePAL whose sequence was only known in the 3' region of the coding region. The similarity among these PAL genes was demonstrated by alignment (Figure 3.12).

It is tempting to suggest that MePAL2 and MePAL3 belong to one gene, and the different base pairs between them may be due to variance between cultivars as they are isolated from different cultivars. MePAL2 was isolated from a genomic library made

from cultivar MBRA 534, whereas MePAL3 was from the PPD-related cDNA library made from cultivar MNGA 1. The hypothesis that they are one gene is also supported by the fact that the differences in nucleic acid sequences of MePAL2 and MePAL3 do not change the amino acid they encode except for one residue (I, isoleucine, to V, valine). At the position of this particular isoleucine residue in the PAL amino acid sequences of other plants, it is either I or V. In order to clarify whether it is true that the difference between MePAL2 and MePAL3 is due to variance between cultivars, Han (2000) amplified and sequenced the PAL gene in cultivar MNGA 1 equivalent to MePAL2 (from MBRA 534), which showed that in the MNGA 1 cultivar there is a PAL gene with exactly the same sequence as MePAL2. This suggests that there is no difference between cultivar MNGA 1 and MBRA 534 in MePAL2 gene. In parsley, PAL1 and PAL2 are of high similarity, only 2 nucleotide difference over 300 bp known sequence by the 5' region of these genes (Logemann *et al.*, 1995). Therefore on the present evidence, in this thesis MePAL2 and MePAL3 are considered different genes (Han, 2000).

```
MePAL1 ATGGAGT---TTTGTGAGGCTCACAATGTAACTGCTTCACCTGGATTCTCCTCTGC------TGA
MePAL2 ATGGCAACAATCTCTCAGAATGGTCACCAGAATGGTTCTTTAGACTCTTTGTGTACAGCTCGTGA
MePAL3 ---------------------------------------------------------------
MePAL  ---------------------------------------------------------------
       ****      * * * **  *    *    * ** ***    *  *  *  * * *      ***

MePAL1 CCCATTGAACTGGGGCATGGCTGCAGAGTCACTGAAGGGCAGCCACCTTGATGAGGTGAA
MePAL2 CCCACTGAATTGGGGTCTGGCTGCTGAGTCTATGAGTGGCAGCCATTTAGACGAAGTGAA
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       **** **** *****  ******* *****  ***  ********  * ** ** *****

MePAL1 GCGCATGGTGGATGAGTACAGGAAGCCTGTGGTGAGGCTAGGGGGTGAGACCCTGACTAT
MePAL2 AAAGATGGTGGCTGAGTTTAGGAAGCCTTTGGTCAAGTTAGGCGGCGAGACCTTGACTGT
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
           ******* *****  ********* **** * * **** ** ****** ***** *

MePAL1 AGCCCAAGTTACAGCAATTGCGAACCATGACTCAGGTGTCAAGGTTGAGCTGTCTGAGGA
MePAL2 AGCTCAGGTGGCAGCTATTGC---TCGTGAATCTGGTCTCCAAGTGGAGCTTGCAGAATC
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       *** ** **  **** *****    * *** ** *** ** * ** *****  * **

MePAL1 GGCTCGAGCTGGGGTCAAGGCCAGTAGTGACTGGGTTCTTGATTCCATGAATAAAGGAAC
MePAL2 TGCTAGAGCTGGTGTTAAGGCGAGTAGTGACTGGGTCATGGATAGTATGAGCAAAGGAAC
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
        *** ******* ** ***** **************  * ***   ****  ********


MePAL1 CGATAGTTATGGTGTCACCACTGGCTTTGGTGCAACCTCCCATAGAAGAACCAAGCAGGG
MePAL2 TGATAGCTATGGTGTCACTACTGGATTTGGAGCCACTTCACATAGAAGAACCAAGCAAGG
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
        ***** *********** ***** ***** ** ** ** ***************** **
```

```
MePAL1 GGGTGCCCTTCAGAGAGAACTCATAAGATTCTTGAATGCTGGGATCTTTGGAAATGGACA
MePAL2 TGGTGCTTTACAGAGGGAGCTCATTAGATTCCTGAATGCTGGGATTTTTGGCAATAAGAC
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
         ***** * ***** ** ***** ****** ************* ***** ***

MePAL1 AGAATCTTGCCACACATTGTCTCACACTGCAACTAGAGCAGCAATGCTGGTGAGAATCAA
MePAL2 AGAATCGTGTCACACTTTGTCACACTCTGCAACAAGAGCAGCAATGCTAGTGAGGATCAA
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       ****** ** ***** ***** *** ******* ************** ***** *****

MePAL1 CACCCTCCTGCAAGGTTATTCAGGCATAAGATTTGAAATCCTGGAAGCCATTACCAAGTT
MePAL2 CACTCTTCTCCAAGGTTACTCAGGCATTAGATTTGAAATCTTGGAAGCTATCACCAAGCT
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       *** ** ** ******** ******** ************ ******* ** ****** *

MePAL1 CATCAACAACAATGTTACTCCGCGTTTGCCCCTCAGAGGCACAATCACAGCCTCTGGTGA
MePAL2 CCTCAACCACAATATTACTCCTTGCTTGCCGCTGAGAGGCACAATCACTGCTTCAGGGGA
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       * ***** ***** *******  * ***** ** ************** ** ** ** **

MePAL1 CCTGGTCCCACTGTCCTACATTGCCGGGCTTTTGACCGGCCGGCCCAACTCCAAGTCGTT
MePAL2 TTTAGTTCCATTGTCCTACATTGCTGGATTGCTCACCGGCCGGCCTAATTCCAAGGCTGT
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
         * ** *** ************** **  *  * *********** ** ****** *  *

MePAL1 AGGGCCCAATGGAGAATCCTTGGATGCAGCTGAAGCCTTTAAGCTTGCTGGGATCAATGG
MePAL2 TGGCCCTAATGGAGAATCCCTAGATGCCCAGCAAGCCTTTCACTCTGCTGGTATTGATTC
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
        ** ** ************ * *****     ******** *   ****** **  **

MePAL1 TGGATTTTTTGAATTGCAGCCCAAAGAGGGTCTAGCTTTAGTAAATGGTACTGCAGTTGG
MePAL2 TGGCTTCTTTGAGTTGCAGCCTAAAGAAGGCCTTGCTCTGGTTAATGGCACTGCTGTTGG
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       *** ** ***** ******** ***** ** ** *** * ** ***** ***** *****

MePAL1 TTCAGGCTTAGCTTCTATGGTTCTTTTTGAGGCCAATGTGTTGGCAGTCCTGTCAGAAGT
MePAL2 TTCTGGCTTGGCTTCCATGGTTCTCTTTGAGGCAAATGTTCTTGCTGTTTTATCAGAAGT
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
       *** ***** ***** ******** ******** ***** * ** ** * ********

MePAL1 TTTATCAGCAATTTTTGCAGAAGTTATGCTTGGAAAACCAGAGTTTACAGATCATTTGAC
MePAL2 CTTATCAGCTATTTTCGCCGAAGTTATGAATGGAAAACCTGAGTTTACTGATCACTTGAC
MePAL3 ------------------------------------------------------------
MePAL  ------------------------------------------------------------
        ******** ***** ** *********  ********* ******** ***** *****

MePAL1 ACATAAATTGAAGCATCATCCTGGACAAATTGAAGCTGCAGCAATTATGGAACATGTCTTG
MePAL2 TCATAAGTTGAAGCACCATCCAGGCCAAATAGAGGCTGCAGCTATAATGGAGCATATTTTA
MePAL3 ------------------------------------------------------------
MePAL  ----------------------------------GCTGCTGCTATTATGGAACACATTTTG
        ***** ******** ***** ** ***** ** ***** ** ** ***** **  * **

MePAL1 GATGGAAGTTCTTATATTAAAGCAGCTCAAAAGGTTCATGAAATTGATCCATTGCAGAAG
MePAL2 GATGGTAGCTCTTATATTAAAGCAGCTAAGAAGTTGCATGAAATTGATCCATTGCAGAAA
MePAL3 ------------------------------------------------------------
MePAL  GATGGAAGTTCATATGTTCAAGAAGCAAAGAAGTTGCATGAGATGGATCCTTTGCAGAAG
       ***** ** ** *** ** *** ***  * *** * ***** ** ***** ********

MePAL1 CCTAAACAAGACAGATATGCTCTTAGAACATCTCCTCAATGGCTTGGCCCTCAGATTGAA
MePAL2 CCAAAGCAAGATCGATATGCTCTCAGAACTTCCCCACAATGGCTAGGTCCTCAGATTGAA
MePAL3 ------------------------------------------------------------
MePAL  CCTAAGCAGGATCGGTATGCTCTTCGTACATCACCACAATGGTTAGGCCCACAAATTGAA
       ** ** ** **  * ********  * ** ** ** ****** * ** ** ** ******
```

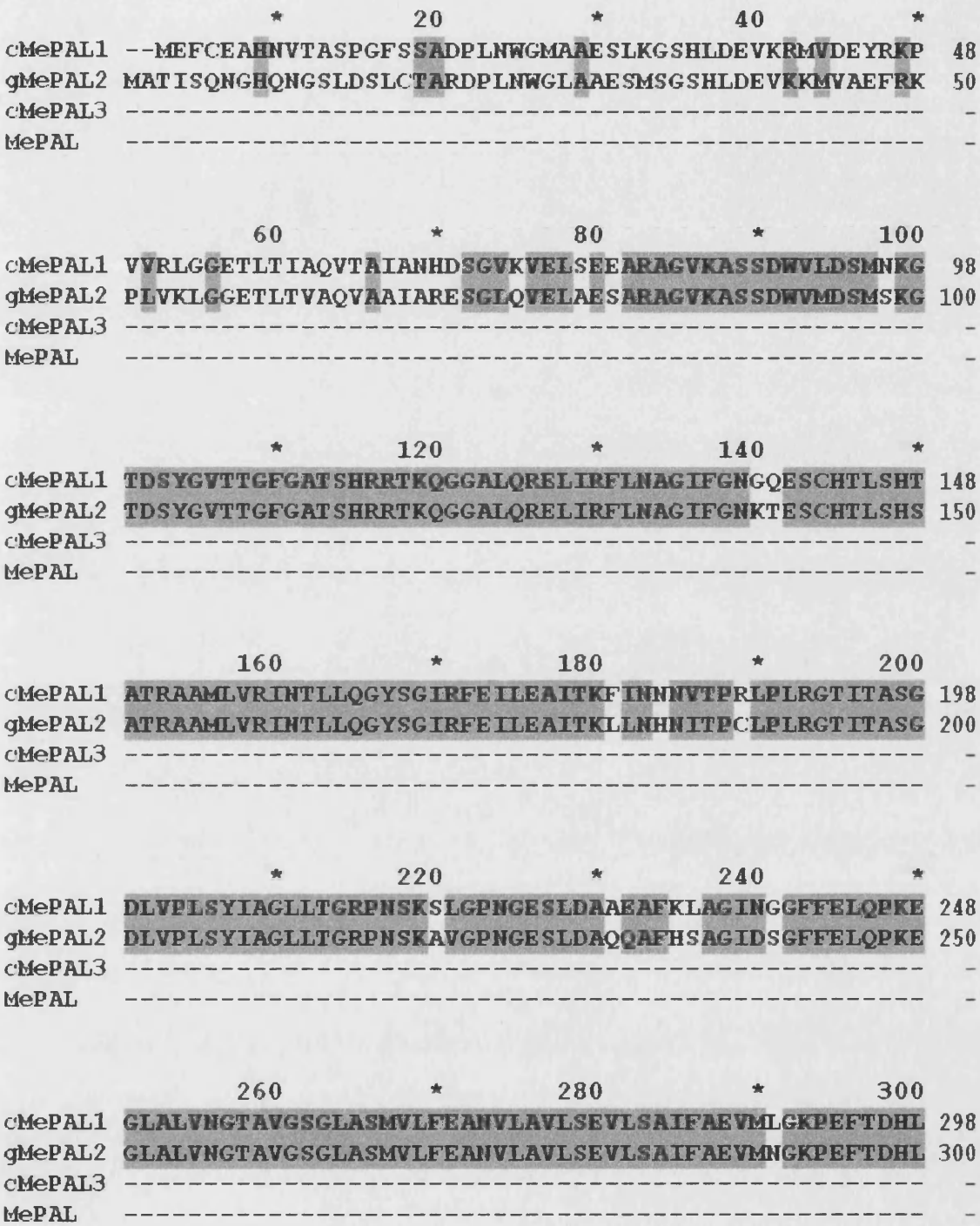

MePAL1 GTGATTCGAACAGCGACTAAAATGATCGAACGCGAAATCAACTCTGTGAATGATAATCCA
MePAL2 GTTATCAGATTCTCCACAAAATCGATCGAAAGAGAGATTAATTCAGTCAATGACAACCCT
MePAL3 ----------
MePAL GTGATTCGATTCTCAACCAAGTCCATTGAAAGAGAGATTAATTCTGTGAATGATAATCCA
MePAL1 TTGATTGATGTCTCCAGGAATATTGCTTTACATGGAGGCAATTTCCAGGGGACCCCAATT
MePAL2 TTGATTGATGTTTCTAGGAACAAGGCCTTGCATGGTGGAAATTTCCAGGGGACCCCAATT
MePAL3 ----------GGGGACCCCAATT
MePAL CTGATTGATGTCTCAAGGAACAAAGCCTTGCATGGAGGAAACTTCCAGGGGACCCCAATT
MePAL1 GGTGTTTCAATGGATAACACTCGTTTAGCCATTGCTTCAATTGGTAAACTCATGTTTGCT
MePAL2 GGAGTCTCAATGGATAATGCACGTTTGGCCATTGCATCAATAGGAAAGCTCATGTTTGCT
MePAL3 GGAGTCTCAATGGATAATGCACGTTTGGCCATTGCATCAATAGGAAAGCTCATGTTTGCT
MePAL GGTGTCTCAATGGACAACACTAGATTGGCATTAGCTTCAATTGGGAAACTCATGTTTGCT
MePAL1 CAATTCTCTGAGCTTGTTAATGATTTTTACAACAATGGGTTGCCTTCAAATCTCACTGGT
MePAL2 CAGTTCAGTGAGCTTGTAAATGATTTTTACAACAATGGGTTGCCATCAAATCTCACAGCC
MePAL3 CAGTTCAGTGAGCTTGTAAATGATTTTTACAACAATGGGTTGCCATCAAATCTCACAGCC
MePAL CAATTCTCTGAGCTTGTTAATGATTTTTACAATAATGGGTTGCCTTCAAATCTCACAGGA
MePAL1 GGACGCAATCCAAGCTTGGATTATGGCTTCAAAGGAGCTGAAATTGCCATGGCATCTTAC
MePAL2 AGCAGGAATCCAAGCTTGGATTATGGCTTCAAGGGAGCTGAAATTGCAATGGCTTCTTAC
MePAL3 AGCAGGAATCCAAGCTTGGATTAcGGCTTCAAGGGAGCTGAAATTGCAATGGCTTCTTAC
MePAL GGAAGGAACCTA-GCTTGGATTATGGTTTCAAGGGTGCTGAGATTGCCATGGCAGCCTAT
MePAL1 TGCTCAGAGCTCCAATTTCTTGCCAATCCTGTAACTAATC-ATGTCCAAAGTGCAGAGCA
MePAL2 TGTTCTGAGCTCCAATACCTTGCAAATCCAGTCACCAGCC-ATGTACAAAGTGCAGAGCA
MePAL3 TGTTCTGAGCTCCAATACCTTGCAAATCCAGTCACCAGCC-ATGTACAAAGTGCAGAGCA
MePAL TGCTCTGAGCTTCAATACCTTGCAAATCCTGTCACAAANCCATGTTCACAGTGCTGAGCA
MePAL1 GCACAACCAAGATGTTAACTCACTAGGCTTGATTTCTTCAAGGAAAACAGCTGAAGCTGT
MePAL2 GCACAATCAAGATGTAAATTCATTGGGGCTAATTTCTTCAAGAAAGACAGAAGAAGCTGT
MePAL3 GCACAATCAAGATGTAAAcTCcTTGGGGCTAATTTCTTCAAGAAAGACAGAAGAAGCTGT
MePAL GCACAACCAAGATG----------
MePAL1 TGACATATTGAAGCTCATGTCTTCTACATACTTAGTTGCTCTATGTCAAGCCATTGACTT
MePAL2 AGACATCTTGAAGCTCATGTCCACGACTTTCTTAGTAGCACTTTGTCAAGCTATTGACTT
MePAL3 AGACATCTTGAAGCTCATGTCCACaACTTTCTTAGTAGCACTTTGcCAAGCTATTGACTT
MePAL ----------
MePAL1 GAGACACTTGGAGGAGAACTTGAAGCAAACAGTCAAGAACACAGTAAGTCAAGTTGCAAA
MePAL2 GAGGCATTTGGAGGAGAACTTGAAGCACGCAGTCAAAAACACAGTAAGCCAAGTAGCTAA
MePAL3 GAGGCATTTGGAGGAGAACTTGAAGCACGCAGTCAAAAACACAGTAAGCCAAGTAGCTAA
MePAL ----------
MePAL1 GAGAGTCTTGACAATGGGCATCAACGGCGAGCTCCACCCGTCGAGATTCTGCGAAAAAGA
MePAL2 GAGGATTCTAACTACAGGAGCTAGTGGAGAACTTCACCCATCAAGATTCTGCGAGAAGGA
MePAL3 GAGGATTCTAACTACAGGAGCTAGTGGAGAACTTCACCCATCAAGATTCTGCGAGAAGGA
MePAL ----------
MePAL1 CCTTCTCAAAGTCGTCGACAGGGAATACGTTTATGCATATGTTGATGATCCTTGCAGTGC
MePAL2 CTTGCTCAAAGTGGTGGATCGCGAGCAAGTCTTCTCTTATGTCGATGACGCCTGCAGTGC
MePAL3 CTTGCTCAAAGTGGTGGATCGCGAGCAAGTtTTCTCTTATGTCGATGACGCCTGCAGTGC
MePAL ----------
MePAL1 AACATACCCATTAATGCAAAAGCTGAGACAAGTACTAGTTGATCATGCCATGATGAATGG
MePAL2 TACCTATCCATTGATGCAAAAACTAAGGCAAGTTCTCGTGGACCATGCCTTGGCAAATGG
MePAL3 TACCTATCCATTGATGCAAAAACTAAGGCAAGTTCTCGTGGACCATGCCTTGGCAAATGG
MePAL ----------

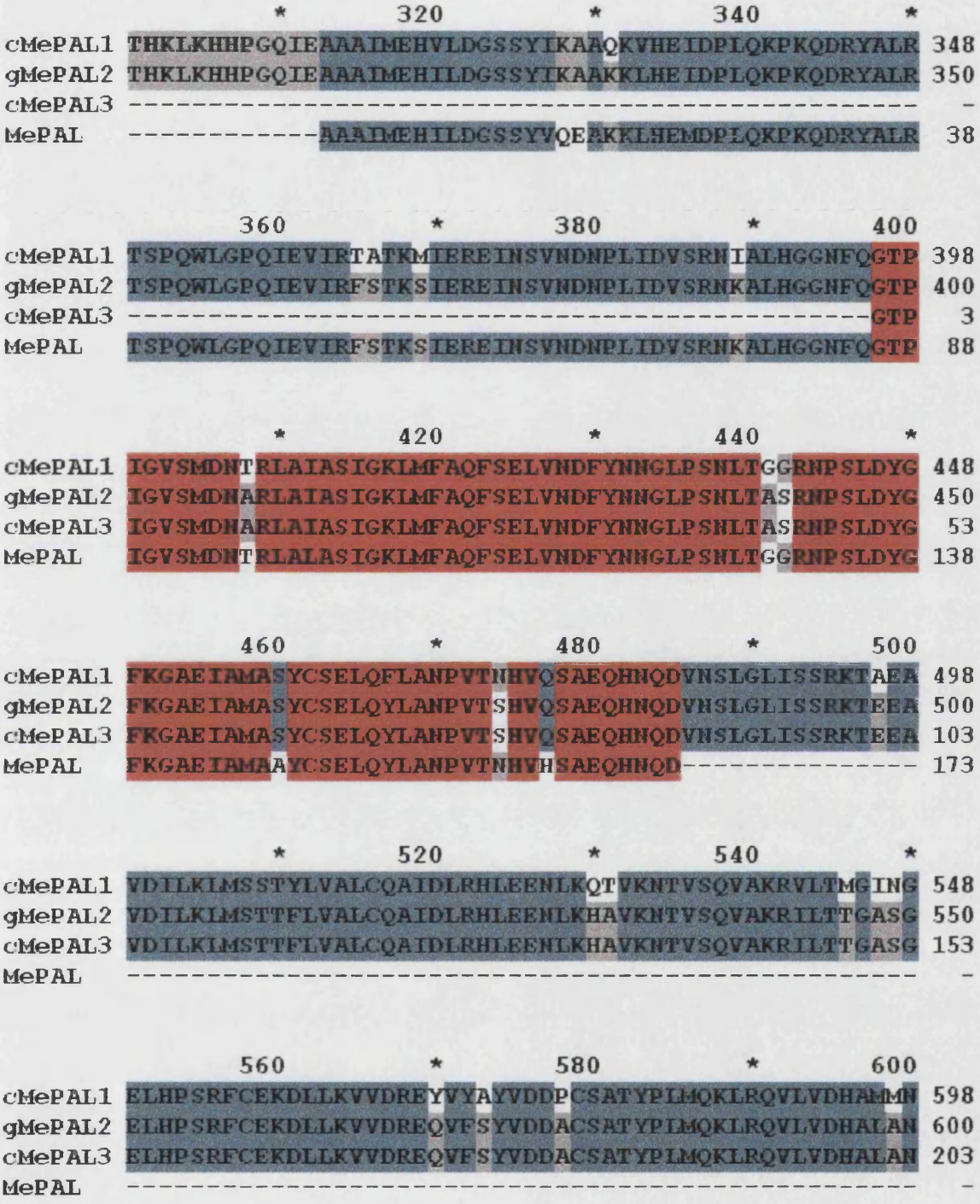
```
MePAL1 TGAAAAGGAGAAGAATTCAAGCACTTCCATTTTCCAAAAAATTGGAGCCTTTGAAGAAGA
MePAL2 CGAGAGTGAGAAGAATGCCAGCACTTCAATCTTCCAAAAGATCAGAGCTTTTGAGGAAGA
MePAL3 CGAGAGTGAGAAGAATGCCAGCACTTCAATCTTCCAAAAGATCAGAGCTTTcGAGGAAGA
MePAL  ------------------------------------------------------------
        ** *  ********* * ******** ** ******** **  **** ** ** *****

MePAL1 ACTCAAGACCCTTTTGCCTAAAGAAGTAGAAAGTGCAAGAACTGAATATGAGAATGGTAA
MePAL2 ATTGAAAGCCCTTTTGCCTAAAGAAGTTGAGAGTGCAAGAGAGGCATACGAGAACGGGAA
MePAL3 ATTGAAAGCCCTTTTGCCTAAAGAAGTTGAGAGTGCAAGAGAGGCATACGAGAAtGGGAA
MePAL  ------------------------------------------------------------
       * * **  ******************* ** *********   * *** ***** ** **

MePAL1 TCCAGCTATTTCTAACAAGATCAAAGAATGTAGGTCATATCCACTATACAAGTTTGTGAG
MePAL2 TCCAGCAATTGCCAACAAGATCAAGGAATGCAGATCTTACCCATTGTATAAGTTTGTGAG
MePAL3 TCCAGCAATTGCCAACAAGATCAAGGAATGCAGATCTTACCCATTGTATAAGTTTGTGAG
MePAL  ------------------------------------------------------------
       ****** *** * *********** ***** ** ** ** *** * ** ***********


MePAL1 GGAAGAACTCGGTTGTAGTTTACTGACCGGCGAGAAGATTCGATCGCCCGGCGAAGAGTT
MePAL2 AGAGGAAATAGGAACTGGGTTGCTCACCGGAGAAAAGATCCGGTCACCGGGAGAGGAATT
MePAL3 AGAGGAAATAGGAACTGGGTTGCTgACCGGcGAAAAGgTCCGGTCACCGGGAGAGGAATT
MePAL  ------------------------------------------------------------
        ** *** * **   * * ** ** ***** ** *** * ** ** ** ** ** ** **

MePAL1 TGATAAGGTTTTCTCAGCAATTTGTGCAGGGAAGCTGATTGATCCCATGCTTGAATGCCT
MePAL2 TGATAAGGTTTTCACTGCTATGTGCCAAGGAAAGATCATTGATCCAATGCTGGATTGTCT
MePAL3 TGATAAGGTTTTCACTGCcATGTGCCAAGGAAAGATCATTGATCCAATGCTGGATTGTCT
MePAL  ------------------------------------------------------------
       ************* * ** ** **   *** *** * ******** ***** ** ** **

MePAL1 CAAGGAGTGGAATGGTGCTCCTCTTCCAATCTGCTAA
MePAL2 CAAAGAGTGGAATGGTGCCCCTCTTCCAATATGTTAA
MePAL3 CAAAGAGTGGAATGGTGCCCCTCTTCCAATATGTTAA
MePAL  ---------------------------------------
       *** ************** *********** ** ***
```

Figure 3.12 Comparison of nucleotide sequences among four PAL genes in coding region.
Nucleotides conserved among MePAL1, MePAL2 and MePAL indicated by *, whilst black and - refer as 'no consensus'. Translation initiation site (ATG) and Termination site (TAA) are in bold. The nucleotides in MePAL3 that are different from MePAL2 are in lower case and bold.

3.8.3.2 Similarity of deduced amino acid sequence

The deduced amino acid sequences of the four cassava PAL genes were compared using multiple alignment (Figure 3.13). MePAL2 shared 93%, 100% and 97% amino acid similarity with MePAL1, MePAL3 and MePAL respectively.


* 20 * 40 *
cMePAL1 --MEFCEAHNVTASPGFSSADPLNWGMAAESLKGSHLDEVKRMVDEYRKP 48
gMePAL2 MATISQNGHQNGSLDSLCTARDPLNWGLAAESMSGSHLDEVKKMVAEFRK 50
cMePAL3 -
MePAL -
60 * 80 * 100
cMePAL1 VVRLGGETLTIAQVTAIANHDSGVKVELSEEARAGVKASSDWVLDSMNKG 98
gMePAL2 PLVKLGGETLTVAQVAAIARESGLQVELAESARAGVKASSDWVMDSMSKG 100
cMePAL3 -
MePAL -
* 120 * 140 *
cMePAL1 TDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNGQESCHTLSHT 148
gMePAL2 TDSYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNKTESCHTLSHS 150
cMePAL3 -
MePAL -
160 * 180 * 200
cMePAL1 ATRAAMLVRINTLLQGYSGIRFEILEAITKFINNNVTPRLPLRGTITASG 198
gMePAL2 ATRAAMLVRINTLLQGYSGIRFEILEAITKLLNHNITPCLPLRGTITASG 200
cMePAL3 -
MePAL -
* 220 * 240 *
cMePAL1 DLVPLSYIAGLLTGRPNSKSLGPNGESLDAAEAFKLAGINGGFFELQPKE 248
gMePAL2 DLVPLSYIAGLLTGRPNSKAVGPNGESLDAQQAFHSAGIDSGFFELQPKE 250
cMePAL3 -
MePAL -
260 * 280 * 300
cMePAL1 GLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEVMLGKPEFTDHL 298
gMePAL2 GLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAIFAEVMNGKPEFTDHL 300
cMePAL3 -
MePAL -


320 340
cMePAL1 THKLKHHPGQIEAAAIMEHVLDGSSYIKAAQKVHEIDPLQKPKQDRYALR 348
gMePAL2 THKLKHHPGQIEAAAIMEHILDGSSYIKAAKKLHEIDPLQKPKQDRYALR 350
cMePAL3 - -
MePAL ----------AAAIMEHILDGSSYVQEAKKLHEMDPLQKPKQDRYALR 38
360 380 400
cMePAL1 TSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNIALHGGNFQGTP 398
gMePAL2 TSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTP 400
cMePAL3 ----------GTP 3
MePAL TSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTP 88
420 440
cMePAL1 IGVSMDNTRLAIASIGKLMFAQFSELVNDFYNNGLPSNLTGGRNPSLDYG 448
gMePAL2 IGVSMDNARLAIASIGKLMFAQFSELVNDFYNNGLPSNLTASRNPSLDYG 450
cMePAL3 IGVSMDNARLAIASIGKLMFAQFSELVNDFYNNGLPSNLTASRNPSLDYG 53
MePAL IGVSMDNTRLALASIGKLMFAQFSELVNDFYNNGLPSNLTGGRNPSLDYG 138
460 480 500
cMePAL1 FKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNSLGLISSRKTAEA 498
gMePAL2 FKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSLGLISSRKTEEA 500
cMePAL3 FKGAEIAMASYCSELQYLANPVTSHVQSAEQHNQDVNSLGLISSRKTEEA 103
MePAL FKGAEIAMAAYCSELQYLANPVTNHVHSAEQHNQD---------- 173
520 540
cMePAL1 VDILKLMSSTYLVALCQAIDLRHLEENLKQTVKNTVSQVAKRVLTMGING 548
gMePAL2 VDILKLMSTTFLVALCQAIDLRHLEENLKHAVKNTVSQVAKRILTTGASG 550
cMePAL3 VDILKLMSTTFLVALCQAIDLRHLEENLKHAVKNTVSQVAKRILTTGASG 153
MePAL - -
560 580 600
cMePAL1 ELHPSRFCEKDLLKVVDREYVYAYVDDPCSATYPLMQKLRQVLVDHALMN 598
gMePAL2 ELHPSRFCEKDLLKVVDREQVFSYVDDACSATYPLMQKLRQVLVDHALAN 600
cMePAL3 ELHPSRFCEKDLLKVVDREQVFSYVDDACSATYPLMQKLRQVLVDHALAN 203
MePAL - -

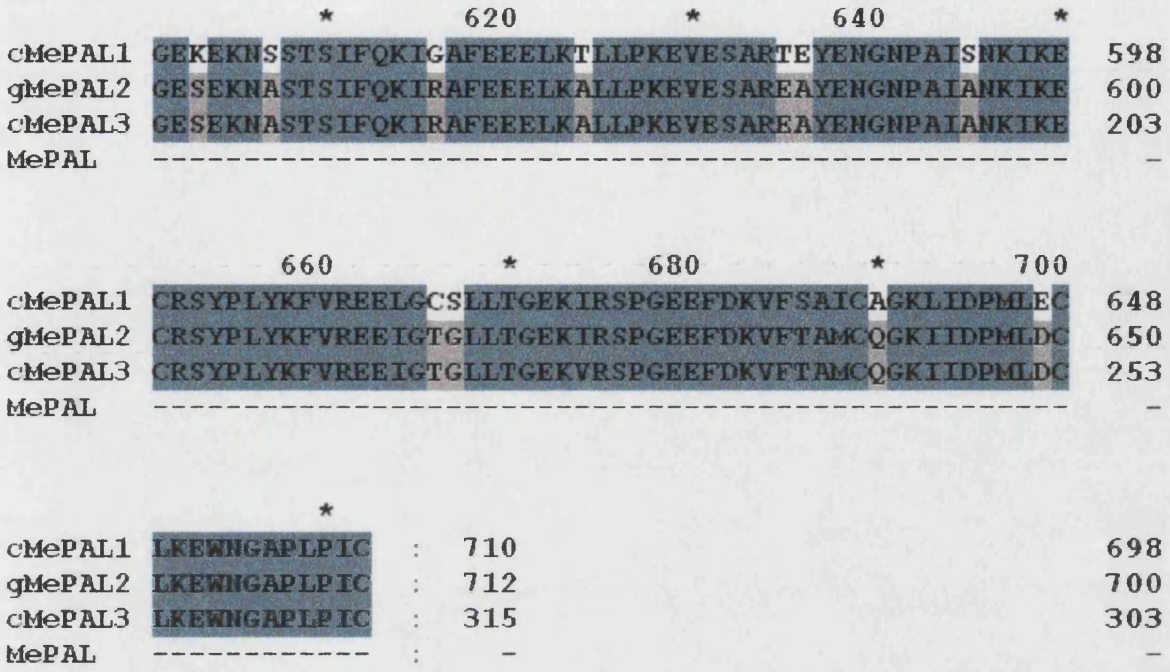


Figure 3.13 **Multiple alignment of the deduced amino acid sequences of four cassava PAL genes.** The residues in dark shading were conserved among all the sequences; those in grey shading conserved among three sequences (the fourth sequence maybe have not always been determined in these regions); those in light grey shading conserved between two sequences. Alignment was made using Genedoc programme at http: //www.psc.edu/biomed/genedoc.

3.8.4 Sequence similarity among PAL genes from cassava and other plants

The coding sequence of MePAL2 was used to search Genebank DNA database by Blastx for amino acid sequence similarity to PAL genes from other plants. The deduced amino acid sequence of MePAL2 showed extensive similarity to those of other cassava PAL sequences, ranging from 70 to 90%. MePAL2 showed 79%, 89%, 92%, 92% and 94% similarity to PAL from rice, *Arabidopsis*, tobacco, lemon and poplar, respectively. To illustrate how the PAL genes were conserved in their amino acid sequence, multiple alignment of amino acid sequences of PAL, including MePAL1, MePAL2, PAL from the tree plants poplar and lemon, dicot tobacco and *Arabidopsis*, and monocot rice and wheat, was carried out. As it can be seen from the alignment (Figure 3.14), these proteins shared most deduced amino acid residues, except for two short stretches at the N-terminal part, one short region near C terminal and a few residues in other regions. The active site, a serine residue in the stretch of sequence GTITSGDLVPLSYIA (Schuster and Retey, 1994, H), was 100% conserved among plant PALs. The sequence around the active site was also highly conserved. Within the sequence of these 15 amino acid residues, the eight PAL sequences in the multiple alignment shared exactly the same residues except one residue in AtPAL2. The similarity of amino acid sequence is even higher in the region encoded by the second exon. MePAL1 & 2 showed higher similarity to dicots than to monocots. It is also interesting to note that MePAL1 & 2 had higher similarity to tree plants poplar and lemon in the deduced amino acid sequence, which may reflect the fact that cassava is more like a tree plant than dicot crops. Similarly, phylogenetic analysis of PAL genes showed that MePAL1 is close to citrus limon. PAL and poplar PALg2b, whereas MePAL2 and MePAL3 are close to poplar PALg1 (Figure 3.15). However, MePAL seems to be close to monocot rice ZB8 and wheat PAL genes.

```
MePAL2          MATISQNGHQ-------NGSLDSLCT----ARDPLNWGLAAESMSGSHLDEVKKMVAEFR
MePAL1          MEFC-EAHN--------VTASPGFSS-----ADPLNWGMAAESLKGSHLDEVKRMVDEYR
PkPALg2b        MEFCQDSRN--------GNGSPGFNT-----NDPLNWGMAAESLKGSHLDEVKRMIEEYR
ClPAL6          MELSHETCNGIKNDRNGGTSSLGLCTG----TDPLNWTVAADSLKGSHLDEVKRMIDEYR
AtPAL2          MDQIEAMLCG-------GGEKTKVAVTTKTLADPLNWGLAADQMKGSHLDEVKKMVEEYR
NtPALB          MASN-GHVN--------GGENFELCKK---SADPLNWEMAAESLRGSHLDEVKKMVSEFR
OsPALZB8        MECENGRVS--------ANGMSGLCVRP--RADPLNWGKATEEMTGSQVDEVKRIGAEYR
TaPAL           MACAWR-------------------SRS--RADPLNWGKAAEELSGSHLEAVKRMVEEYR
                *                          *****  *::.: **::: **::  *:*
```


MePAL2 KPLVKLGGETLTVAQVAAIARES----GLQVELAESARAGVKASSDWVMDSMSKGTDSYG
MePAL1 KPVVRLGGETLTIAQVTAIANHD---SGVKVELSEEARAGVKASSDWVLDSMNKGTDSYG
PkPALg2b NPVVKLGGETLTIGQVTAIASRD---VGVMVELSEEARAGVKASSDWVMDSMSKGTDSYG
ClPAL6 RPVVKLGGESLTIGQVTAIAAHD---SGVKVELAEAARAGVKASSDWVMDSMMKGTDSYG
AtPAL2 RPVVNLGGETLTIGQVAAISTVG---GSVKVELAETSRAGVKASSDWVMESMNKGTDSYG
NtPALB KPMVKLGGESLTVAQVAAIAVRDKSANGVKVELSEEARAGVKASSDWVMDSMNKGTDSYG
OsPALZB8 QPLVKIEGASLRIAQVAAVAAAG----EARVELDESARERVKASSDWVMNSMMNGTDSYG
TaPAL KPVVTMEGATT-IAMVAAVAAGS----DTRVELDESARGRVKESSDWVMNSMMNGTDSYG
MePAL2 VTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNKTESCHTLSHSATRAAMLVRINTLLQ
MePAL1 VTTGFGATSHRRTKQGGALQRELIRFLNAGIFGNGQESCHTLSHTATRAAMLVRINTLLQ
PkPALg2b VTTGFGATSHRRTKQGGELQKELIRFLNAGIFGNGTESSHTLPRSATRAAMLVRTNTLLQ
ClPAL6 VTTGFGATSHRRTKQGGALQKELIRFLNSGIFGNGTESSHTLPHSATRAAMLVRVNTLLQ
AtPAL2 VTTGFGATSHRRTKNGTALQTELIRFLNAGIFGNTKETCHTLPQSATRAAMLVRVNTLLQ
NtPALB VTTGFGATSHRRTKNGGALQKELIRFLNAGVFGNGTETSHTLPHSATRAAMLVRINTLLQ
OsPALZB8 VTTGFGATSHRRTKEGGALQRELIRFLNAGAFGTGTDG-HVLPAEATRAAMLVRINTLLQ
TaPAL VTTGFGATSHRRTKEGGALQRELIRFLNAGAFGTGTDG-HVLPAAATRAAMLVRVNTLLQ
MePAL2 GYSGIRFEILEAITKLLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPNG
MePAL1 GYSGIRFEILEAITKFINNNVTPRLPLRGTITASGDLVPLSYIAGLLTGRPNSKSLGPNG
PkPALg2b GYSGIRFEMLEAITKMINHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPNG
ClPAL6 GYSGIRFEILETITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGSNG
AtPAL2 GYSGIRFEILEAITSLLNHNISPSLPLRGTITASGHLVPLSYIAGLLTGRPNSKATGPDG
NtPALB GYSGIRFEILEAITKLINSNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPNG
OsPALZB8 GYSGIRFEILEAIAKLLNANVTPCLPLRGTITASGDLVPLSYIAGLVTGRENAWALPPDG
TaPAL GYSGIRFEILETIATLLNANVTPCLPLRGTITASGDLVPLSYIAGLVTGRPNSMATAPDG
MePAL2 ESLDAQQAFHSAGIDSGFFELQPKEGLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAI
MePAL1 ESLDAAEAFKLAGINGGFFELQPKEGLALVNGTAVGSGLASMVLFEANVLAVLSEVLSAI
PkPALg2b EPLTPAEAFTQAGIDGGFFELQPKEGLALVNGTAVGSGLASMVLFEANVLAILSEVLSAI
ClPAL6 QVLNPTEAFNLAGVTSGFFELQPKEGLALVNGTAVGSGLAATVLFEANILAIMSEVLSAI
AtPAL2 ESLTEKEAFEKAGISTGFFDLQPKEGLALVNGTAVGSGMASMVLFEANVQAVLAEVLSAI
NtPALB ETLNAEEAFRVAGVNGGFFELQPKEGLALVNGTAVGSGMASMVLFDSNILAVMSEVLSAI
OsPALZB8 SKVNAAEGLKIAGIQGGFFELQPKEGLAMVNGTAVGSGLASTVLFEANILAILAEVLSAV
TaPAL SKVNAAEAFKIAGIQHGFFELQPKEGLAMVNGTAVGSGLASMVLFEANVLSLLAEVLSGV
MePAL2 FAEVMNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYIKAAKKLHEIDPLQKPKQD
MePAL1 FAEVMLGKPEFTDHLTHKLKHHPGQIEAAAIMEHVLDGSSYIKAAQKVHEIDPLQKPKQD
PkPALg2b FAEVMQGKPEFTDHLTHKLKHHPGQIVAAAIMEHILDGSAYVKEAQKLHEIDPLQKPKQD
ClPAL6 FAEVMNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAQKLHETDPLQKPKQD
AtPAL2 FAEVMSGKPEFTDHLTHRLKHHPGQIEAAAIMEHILDGSSYMKLAQKVHEMDPLQKPKQD
NtPALB FAEVMNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAQKLHEMDPLQKPKQD
OsPALZB8 FCEVMNGKPEYTDHLTHKLKHHPGQIEAAAIMKHILEGSSYMKHAKKLGELDPLMKPKQD
TaPAL FCEVMNGKPEFTDHLTHKLKHHPGQIEAAAIMEHILEGSSYMMLAKKLGELDPLMKPKQD
MePAL2 RYALRTSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSM
MePAL1 RYALRTSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNIALHGGNFQGTPIGVSM
PkPALg2b RHALRTSPQWLGPLIEVIRTSTKMIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSM
ClPAL6 RYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSM
AtPAL2 RYALRTSPQWLGPQIEVIRQATKSIEREINSVNDNPLIDVSRNKAIHGGNFQGTPIGVSM
NtPALB RYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSM
OsPALZB8 RYALRTSPQWLGPQIEVIRAATKSIEREINSVNDNPLIDVSRGKALHGGNFQGTPIGLSM
TaPAL RYALRTSPQWLGPQIEVIRAATKSIEREINSVNDNPLIDVSRGKAIHGGNFQGTPIGVSM

Figure 3.14 Multiple alignment of deduced amino acid sequences of PAL genes from cassava and other plants.

The sequence around the active site (serine) of PAL is in bold. The junction between exon1 and exon2 is indicated by an arrow in red. * - single, fully conserved residue; : - conservation of strong groups; . - conservation of weak groups; - no consensus

Me: *Manihot esculenta* (cassava); Pk: *Populus kitakamiensis* (poplar) (Access No. D43802); At: *Arabidopsis thaliana* (L33678); Nt: *Nicotiana tabacum* (tobacco) (D17467); Os: *Oryza sativa* (rice) (Zhu *et al.*, 1995); Ta: *Triticum aestivum* (wheat) (X99705).

3.8.5 Introns of PAL genes

The introns of similar genes are usually quite different in their sequences. Blast searches using the sequence of the MePAL2 intron revealed no similarity with any sequence in Genbank. However, comparison between the introns of MePAL1 and MePAL2 showed that there was 68% similarity over a region of 140 bp. Within this region, there were three stretches of sequences shared between the introns from these two genes, which are ATATATTAA, TAAATGATA and ATTCATA. Like the cassava PAL genes described here, most PAL genes from plants contain only one intron (Table 3.1). But there are a few exceptions. In pine trees (*P taeda* or *P. banksiana*), the only gymnosperm in which PAL genes were studied, there is no intron in the PAL genes. In *Arabidopsis*, PAL3 differs significantly from PAL1 and PAL2 and other sequenced plant PAL genes in that it contains an additional intron and its deduced amino acid sequence is less similarity to other PAL proteins. It is interesting to note that a PAL gene in the fungus *Rhodosporidium toruloides* contains five introns (Rasmussen and Oerum, 1991). The position of the intron is also conserved among most PAL genes, it was located about one-fifth into the coding region, or around RF residues in QRELI**RF**LNAGN (Figure 3.14). If the position of the intron is marked from the 3' end of the second exon, many PAL genes would share the same position of intron. As shown in Table 3.1 the sizes of the second exon of 14 PAL genes ranged from 1739 to 1752, with 11 of them being 1747, 1748 and 1749. Although the number and position of intron(s) are conserved among PAL genes, the size of the introns of PAL genes varies greatly, ranging from 177 to 1972 bp.

Table 3.3 Size comparison of exons, intron and putative peptide of PAL genes from different plant species

Genes	Exon1	Exon2	Amino acid	Intron legth	Reference
MePAL2	388	1748	712	711	This thesis
MePAL1	381	1748	710	598	Han, 2000
AtPAL1	427	1748	725	449	Wanner, 1995(23)
AtPAL2	404	1748	717	209	Same as above
NtPAL1	398	1748	715	1932	Fukasava,1996(50)
NtPALA	398	1749	716	1972	AB008199, Genebank
NtPALB	397	1749	715	1952	AB008200, Genebank
TsPAL	428	1747	725	616	Howles , 1994
OsPALZB8	391	1742	710	177	Zhu *et al.*, 1995
OsPALGP1	367	1739	701	1310	Minami *et al.*,1989
PsPAL1	424	1747	726	1551	D10002, Genebank
TaPAL	355	1747	701	118	X99705, Genebank
PkPALg2b	386	1747	710	817	Osakabe Y,1995
PkPALg1	394	1752	715	884	D30656, Genebank

Me: *Manihot esculenta* (cassava) At: *Arabidopsis thaliana* Nt: *Nicotiana tabacum* Os: (rice) *Oryza sativa* Ta: *Triticum aestivum*
Ps: (Pea) *Pisum sativum* PK: (poplar) *Populus kitakamiensis* Potato: *Solanum tuberosum* Ts: *Trifolium subterraneum*

3.9 Summary

Like PAL genes in other plant systems, PAL was encoded by a multigene family in cassava, estimated to be of at least 4 genes using Southern analysis. One PAL gene, MePAL, had been partially sequenced (Pereira et al., 1999); another two, cDNA MePAL1 and part of cDNA MePAL3 were isolated by Han (2000) from a PPD-related cDNA library. In order to isolate other PAL genes and the genomic correspondent of MePAL1 and MePAL3, a cassava genomic library was screened with probe made from 1.9 kb of cDNA MePAL1 fragment, resulting in the isolation of MePAL2. MePAL2 had two exons and one intron, 388 bp, 1748 bp and 711 bp respectively. 2741 bp 5' and 960 bp 3' flanking regions of MePAL2 were also sequenced. MePAL2 is similar not only to other cassava PAL genes, but also to PAL genes from other plants, in nucleotide and amino acid sequence, and also in structure such as the number and position of intron. A genomic clone corresponding to cDNA MePAL1 could not be isolated from the genomic library despite employing several different strategies. However, its genomic equivalent was successfully amplified from genomic DNA by PCR and sequenced. Cassava PAL genes shared higher similarity to amino acid sequences with from other plants especially poplar and lemon.

CHAPTER FOUR: CHARACTERIZATION OF THE MEPAL2 GENE PROMOTER

4.1 Introduction

The promoter, as a control region of a gene, is usually located upstream to the coding sequence of a gene or an operon, and is the critical component for the regulation of gene expression, in organisms from *E.coli* to higher plants and mammals. It binds and directs RNA polymerase to the correct transcriptional start site and thus permits the initiation of transcription. The analysis of gene promoters involves two main areas. One of them is the identification of potential regulatory regions (*cis*-acting sequences) within gene promoters. This is normally achieved by constructing a deletion series of promoter fragments, which are fused with a reporter gene whose level of expression is then analysed *in vivo*. Secondly, following on from deletion analysis, *trans*-acting factors are identified that bind to regions of promoters and facilitate gene transcription. There has been a rapid progress in both of these areas over the last twenty years and these approaches have proved fruitful as starting points to identify factors controlling gene expression. As phenylalanine ammonia-lyase (PAL) catalyses the core reaction of phenylpropanoid metabolism, its gene promoters have been extensively studied in a number of higher plants, such as the dicots *Arabidopsis*, bean, parsley, and the monocot, rice (Ohl *et al.*, 1990; Bevan *et al.*, 1989; Liang *et al.*, 1989; Schulze-Lefert *et al.*, 1989; Zhu *et al.*, 1995). The regulatory properties of PAL genes during plant development and in response to stress treatments have been analysed (Dixon *et al.*, 1995; Liang *et al.*, 1989; Lois and Hahlbrock, 1992; Reinold *et al.*, 1996; Subramaniam *et al.*, 1993; Wu and Hahlbrock, 1992). Experiments using transgenic plants expressing a β-glucuronidase (GUS) reporter gene under the control of PAL promoters have shown that transcriptional activities of these PAL genes display differential spatial and temporal patterns during development and response to environmental cues (Bevan *et al.*, 1989; Kawamata *et al.*, 1997; Liang *et al.*, 1989; Mauch-Mani and Slusarenko,

1996; Ohl *et al.*, 1990; Shufflebottom *et al.*, 1993). P, A and L boxes or AC-rich *cis*-acting elements have been identified within PAL promoters and the promoters of genes encoding other phenylpropanoid enzymes such as cinnamate-4-hydroxylase (C4H) and 4-coumarate:CoA ligase (4CL) (Feldbruegge *et al.*, 1997; Grotewold *et al.*,1994; Sablowski *et al.*, 1994). These elements serve as binding sites for transcription factors with myb DNA binding domains and play central roles in the control of developmentally regulated PAL expression (Hatton *et al.*, 1995; Hatton *et al.*, 1996; Leyva *et al.*, 1992).

In this chapter, the promoter sequence of the MePAL2 gene was analysed for putative *cis*-acting elements by comparison with the promoters of PAL genes and related genes from other plants. A MePAL2 promoter deletion series and MePAL2 promoters with mutated TATA boxes were made and fused with the GUS gene. The promoter activities in these constructs were analysed by transient assay using cassava embryogenic suspension cell culture and transformation systems.

4.2 Identification of putative regulatory sequences within the 5'-flanking sequence of the MePAL2 gene

To identify putative regulatory sequences within the promoter of the MePAL2 gene, the promoter was analysed by comparison with other gene sequences or *cis*-elements or motifs of other promoters. Blast searches using the MePAL2 promoter region including the initial 913 bp and extended 2740 bp sequence revealed no similarity with any sequences in Genebank. Every 50 bp of the promoter were used in Blast searches in order to find genes containing domains similar to MePAL2, but no similarity was found over the stretches of 50 bp. A large number of *cis*-acting sequences have been identified in plants. They mainly consist of short stretches of DNA, usually less than 20 bp in length, which have been reported to play a role in regulating the transcription of the gene in which they are contained. In order to analyse further the structure and decipher the possible expression pattern of MePAL2 promoter, the 5'-flanking sequences were searched for the presence of such sequence motifs from related genes. Several motifs or near-match motifs were found to be present throughout the length of the 5'flanking sequences. They were the TATA box, CCAAT box, G box, P, A, L boxes and several other motifs. The results of the analysis are presented in Table 4.1 and Figure 4.1.

TATA boxes, which are important protein binding motifs in a promoter, have been found to be the most common motif in both prokaryotic and eukaryotic gene promoters though the positions are different in prokaryotic and eukaryotic promoters. In plants, the TATA box is close to the transcriptional start site (TS), and is normally located about 16 to 54 bp 5'of the TS (Hughes, 1996). It is involved in orientation of RNA polymerase II, the enzyme responsible for mRNA synthesis (Hughes, 1996). Upstream from the putative translation start codon in MePAL2 promoter, there are two TATA-like boxes TATAT (-187 to -183) and TATTTAAA (-144 to -137). This raises the question as to which TATA box is the functional protein binding site.

Motif	Sequence in gMePAL2	Location in gMePAL2	Reference
TATA box	1. TATA 2. TATTTAAA	-187 to -183 -144 to -137	Watson *et al.*, 1992
CAAT box	CCAAT	-990 to -986 -839 to -834 -402 to -399	Watson *et al.*, 1992
P box, (C/T)T(T/C)(C/T)(A/C)(A/C)C(A/C)A(A/C)C(C/A)(C/A)C	TTTCAACCAACCCA	-248 to -235	(Logemann *et al.*, 1995)
A box, CCGTCC L box, (T/C)C(T/C)(C/T)ACC(T/A)ACC	CCCTCC GCTTACCTACC	-219 to -214 -209 to -199	
AC-I AC-II, CCACCAACCCC	CCCACCTACC ACCAACCA CAACCAACCC ACCTACCA	-1236 to –1227 -336 to –359 -245 to –236 -1233 to -1226	(Gray-Mitsumune *et al.*, 1999; Hatton *et al.*, 1995)
Box I, (T/A)CT(C/A)ACCTA(C/A)C(C/A) Box II, CCA(A/C)C(A/T)AAC(C/T)CC Box IV, TAATTAA Box V, ATTAGTAAGTGAT	CTTACCTACCCA CAACCAACC TAATTAA ATTAGTAAGTGAT	-208 to -199 -246 to -238 -2170 to -2164 -2018 to -2012 -1125 to –1119 -437 to –426	(Ohl *et al.*, 1990)
Jasmonic acid responsive, TGGAGGCTGGTTGGAGA	AGGTGGCTGTAGCGTTGGAGA	-425 to -405	(Pastuglia *et al.*, 1997)
H-box	CCTACC	-1232 to -1227 -204 to -199	(Loake *et al.*, 1992)
G-box	CAGGTG CACGTG	-2291 to -2286 -1658 to -1653	(Loake *et al.*, 1992)
Auxin inducible	AATAAG	-2204 to -2199	(Liu *et al.*, 1994)
Ethylene inducible	AGCCGCC	-2529 to -2524	(Ohme-Takagi *et al.*, 2000)
Direct repeats	AATTAGC(AAATTCGGTT)AATTAGC	-753 to -730	(Seki *et al.*, 1996)
Inverted repeats	1. ACTCTTAC(CAATA)CATTCTCA 2. GCACCACCACCACG 3. TCTCTTCA(CCTCT)ACTTCTCT 4. TTCCTCT(CTCAGATCCTTTGTT)TCTCCTT	-846 to –826 -346 to –339 -177 to –164 -85 to -63	(Seki *et al.*, 1996)
AT domains	1. TATTTATTTTAATAATAATAATAATTTTTA 2. AAAAATATTTAATTATATAATATAAAATTAAAAT	-901 to –871 -683 to -650	

Table 4.1 DNA sequence motifs located within the 5'-flanking region of MePAL2

Motifs that have been shown to have a regulatory role in other plant gene promoters were identified in the promoter of MePAL2. Some of the motifs were selected from genes that are either phenylpropanoid related gene or stress-related. Numbering of the motifs within the MePAL2 5'-flanking sequence is with respect to the translation start site.

-2740 CTTCGTTTCC

-2730 CACAGACGGCGCCAATTGATGATCTGAGATCCAATAGATTAGGGTTTTACAAGGGTTTTG

-2670 TAATGAATAATGAAAGCTTAAACTTCCTGGGGAGGGGACTCCCCTTTTATACATTGTCTT

-2610 GCTTGCTGGTGACGTGTAAAGGCCTCGCCATGATTGGGACACGCGTCCCTGACATACAGA

ethylene inducible

-2550 TTCGGGCGTACGAGGGTATCAGCCGCCTGCTCCATGCGTAACGGCCTCTGATTCCCCCCT

-2490 GGGCGTACGTTTGTGCGGACCTGCTAGGCTGTCTGGTGAGTACCATTCTGGGTTCTGGGC

-2430 CGGACTGGAGAGTGGGCCTGATGTTGGGCCTGGGTGGAGAGGATCTGCTCGAGGGAGTAT

-2370 TCGGGCTGGACCATCCTGGGTGAGAAGTCGTAGCCGAGCCCGGCCTCCGGCGCTGAGGGA

G box

-2310 GCTGGACCTGCTTCATGGCCAGGTGCATAGGCCTTCTCTCCATGGGCTTTTCGTTGTGGG

auxin inducible

-2250 CTAGGCCCACGATGGGGATAAGGAAATCCAGCGGTCATCAAAAATTAATAAGAGAAATAT

box IV

-2190 AAAGAAAGGATAAGATTGGTAATTAAGCCAAATGAATACAAAGGTTTCATGACAGAGAAT

-2130 AAAGGTTTCTAACTCCTTTTTAAAAAACAAAATTCATTATGTAATCTTCTCTCTTTTGTG

box IV

-2070 TAAGGTTTCATAGTAAAAATAAAAATTTATTTCATAAAATTCATTATATAAATAATTAAG

-2010 TAAAATTATTAATCTATAAAAATTTATAATACCTAAGTAAAATTAAGAATATAATTACTG

-1950 ATAACCGTCGGATATTCATTTTATTGGATCAGGGCGAATCCAAAAAAGTAGAATCAGTTC

-1890 AAAACTTAGTAGATACTCTAGAATCACCCAATCCGACCCTTTTAAGTTGGGCTCGAGGGC

-1830 ATGAAAAGGGAGTCAGATCGCATATGTGCTTTAACTCAGCATGTTAAAGATCGACTCAGT

-1770 GACGTGAAAAGAAGTGAGATAGCAGATCCGGTCACTTTATAACCGTTTTCGTTTATACGC

G box

-1710 CAAAAAATATTAAATGACCGCCTGATACGGAATGAATCTCTGGATATCCGTACACGTGGG

-1650 TCAGTGTGTCAGGACAAATGGTATTATAACAGAAAGACCATTACACATTTAAATAAAAAG


-1590 AAAAAAAAATAAGACGAAAAAAGAATTATCTAGAAGGACCAAATTTACACATACATACTA
-1530 AAAATCTTTTTATTCTCTTAGTCAGAACATCAATTACTTTTTATAAATTTAGAGATTAAA
-1470 TAGTTTATTATAATTTAATCAGTAATTTTTTTTTATATATTATTTTAAAGCAATATTTGG
-1410 TTAGTGTTTATAATATATTGACCATTATTTTGGCTTCTTATTACAAGTAGAGCCATTGCT
-1350 AGCTAATGACAATATTCTATTATTAATGGTCAACAAAGGTAAGCTTaAAGCTAAAGTCCT
AC I
-1290 TAATCAATTGCTCAGATTTTATCAAAATAACTTCAATTCTTAAGAGATACTATTCCCACC
H box
-1230 TACCATCACTCTATTTTTCTTTCTTTTCTTTTTTCTTTTTTAGATTTTTAATCACTGCCA
-1170 CCTAAACTATGTAATCTCTCATGGAAAATATTTTTTAACACATAAAAAATAAAAAAATAG
-1110 AAAATTACACATTCATTATAATCTAAATATCAGTAAAACATAGAAAATTTAATTACTGAA
box IV
-1050 ATTTTATTTAACATAGAATAAAATTTAATTAATTTAGTGATGTTTCTCACTTCTCAGCCT
CCAAT box
-990 CCAATGGCAAAGATCGATCCAAAGAAATCATAGCTGTCCATGTTAGGGTTCCTCAAATCT
AT-rich domain
-930 CGTGTATATTTTAAACTAAGCTTATTGTGTTATTTATTTTAATAATAATAATAATTTTTA
CCAAT box inverted repeat
-870 CCCAAAAAAACAACAACATATTTAACTCTTACCAATACATTCTCACTTTATATTTATATT
-810 TATAACAAAATACTATAAGAAATTTTAATAAAAATAAAAGATAAAACTATAGACTAAAAT
direct repeats
-750 TAGCAAATTCGGTTAATTAGCTTCAATTTTTTAGTTATGTACAACACACAATTAATATTT
AT-rich domain
-690 TAATTACAAAAATATTTAATTATATAATATAAAATTAAAATCTCATTGTTTAAATTTTTT
-630 ATTATATGTTAAAAATTACATATTTTTATTTAATAAAAGGGTTTATATTTATATATTTTT
-570 TTATAAAATTATGTCATATTTATACTAAATTAATTTATTAAAAATTAAAATACTTTTAAT
-510 GAATATTTATTTTATTTCTTCAAAAAAAAAGTTTCAATTTTACTTTAAAAAAAATCAAAA
box V, JA responsive element CAAT box
-450 TTTATTTTGAAATTTACTTACTGATTAGGTGGCTGTAGCGTTGGAGACCCAATCAATGGA

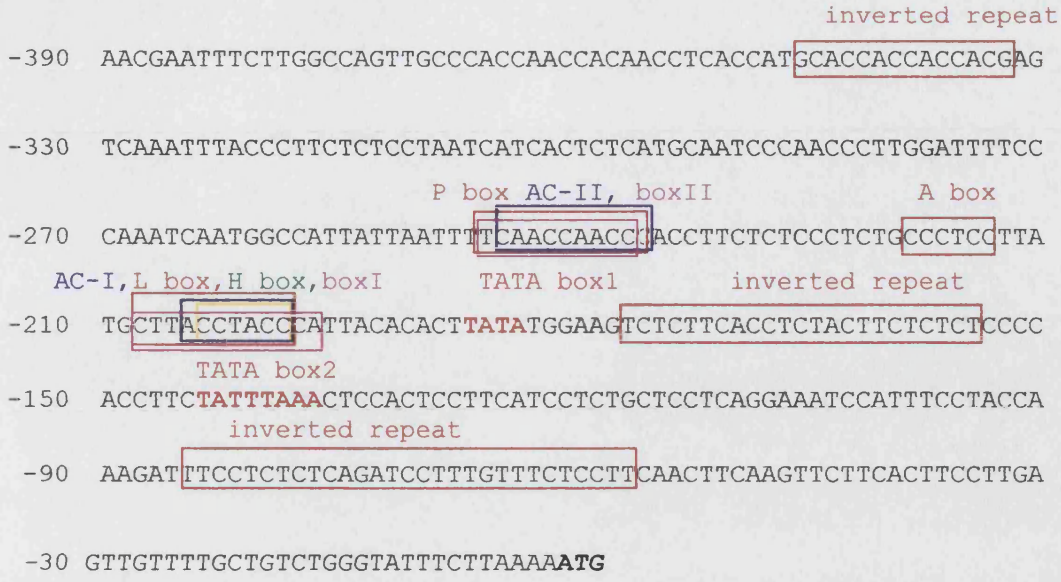


Figure 4.1 Spatial distribution of putative *cis*-acting elements within the MePAL2 5'-flanking sequences.

Putative motifs are in boxes. Putative TATA boxes are in bold red. The sequence is numbered with translation start site ATG (in bold) as +1.

Another sequence motif commonly found within eukaryotic gene promoters is the CCAAT box. Its position varies in different promoters but it is always located further from the gene than is the TATA box. In MePAL2 promoter, three CAAT box-like motifs were found and they were located at -990 to –986, -838 to –834 and –402 to –399 respectively (Figure 4.1). The function of CAAT *cis*-acting element has been tested in several plants. The functional map of a nopaline synthase (*nos*) promoter showed that the "CAAT" box is essential for maximal activity of the promoter. Deletion of this sequence reduced apparent nos expression by over 80% (Shaw *et al*, 1984). Furthermore, the responses of the nopaline synthase (*nos*) promoter to H_2O_2, methyl jasmonate and wounding in tobacco were significantly reduced by deletions of the CAAT box region and the sequence between –112 and –101 (Dai and An, 1995).

G-boxes, CACGTG or CAGGTG, were found at –2291 bp to –2286 bp and –1658 bp to -1653 bp of the MePAL2 promoter, respectively (see Figure 4.1 and table 4.1). The G-box has been identified in a range of plant gene promoters and shown to play a role in the regulation of gene expression by developmental cues, ABA, light, UV irradiation and in response to stresses. A family of bZIP proteins that bind G-box sequences has been described (Foster *et al*, 1994; Menkens *et al.*, 1995). Particularly, the G-box seems to be involved in the control of gene expression upon wounding. In the extensin *extA* promoter of *Brassica napus*, a motif (GT**ACGTGT**TATAAA**ACGTGT**) containing two G-box direct repeats, controlled the induction of the *extA* gene expression in response to wounding (Elliott and Shirsat, 1998). The promoter of a horse-radish peroxidase prxC2 gene failed to respond to wounding after a G-box-like motif CACGTG was deleted from the promoter (Kawaka *et al.*, 1994). Mutation analyses of the bean seed storage protein phaseolin promoter demonstrated that the CACCTG motifs act as positive *cis*-elements (Kawagoe *et al.*, 1994).

H-boxes, CCTACC, another common stress inducible *cis*-element in plant promoters such as PAL, 4CL and chalcone synthase (*chs*) (Lois *et al*, 1989; *Ohl et al.*, 1990), were found at –1232 bp to –1227 bp and –204 bp to –199 bp of MePAL2 promoter. The G-box (CACGTG) and H-box (CCTACC) *cis* elements function in the activation of phenylpropanoid biosynthetic genes involved in the elaboration of lignin precursors, phytoalexins and the secondary signal, salicylic acid, as early responses to pathogen

attack (Droge-Laser *et al.*, 1997). H-box elements contained within the bean chalcone synthase gene *chs* 15 promoter, along with a G-box, have been shown to be important for regulating both the developmental expression of the gene in root tips and corolla and the induction of the gene in elicited protoplasts and transgenic tobacco (Yu *et al.*, 1993). Further evidence of regulatory role for the H-box element came from the purification of two proteins, KAP1 and KAP2, from bean cell suspension culture that bind the *chs*15 H-boxes (Yu *et al.*, 1993). A novel bZIP protein, G/HBF-1, was identified to bind to both the G-box and adjacent H-box in the proximal region of the chalcone synthase *chs15* promoter (Droge-Laser *et al.*, 1997).

Three *cis*-acting elements, designated as P, A and L boxes, have been identified and shown to be conserved among phenylpropanoid genes, such as PAL and 4CL, in many plants (Table 4.2). These elements are AC-rich and were therefore also termed AC-I (L box) and AC-II boxes (P box) (Sablowski *et al.*, 1995). In PsPAL1 of pea, the L and P boxes were named Box I and Box II (Ohl *et al.*, 1990) (Figure 4.1 and Table 4.1). These elements have been shown to be inducible by elicitor and light *in vivo* (Lois *et al.*, 1989). In bean (*Phaseolus vulgaris*), *in vitro* footprinting has identified that the AC-I and AC-II *cis*-elements, plus AC-III (another AC rich *cis*-element) and G box were involved in determining tissue-specific expression of the bean PAL2 promoter. All four elements were involved in specific expression in stem cortical cells and the pigmented regions of petals. The three AC elements were also involved in conferring expression in xylem tissue. The AC-I element was essential for root tip expression, whilst the AC-II and G-box elements also contributed to the expression in root tips (Hatton *et al.*, 1995). Comparison of the MePAL2 promoter sequence with the sequences of these elements showed that there were such elements in the promoter. P, A and L boxes were located at –248 to –235, -219 to -214 and –209 to -199. The order and relative positions of these boxes were different in the promoters of different plants. In promoters of parsley (*Petroselimum crispum*) PcPAL (1, 2 & 3) and potato (*Solanum tuberosum*) StPAL (1 & 2), there were arranged as P-A-L, the same order as in the cassava MePAL2 promoter (Logemann *et al.*, 1995; Lois, *et al* 1989; Joos *et al.*, 1992). But the distances between the P, A and L boxes were different. The distance between P and A was 15 to 47 bp, and 53 to 68 bp between A and L in PcPALs and StPALs. In the cassava PAL promoter, these were 28 bp between P and A, 11 bp between A and L. In

some PAL promoters the P, A and L boxes were in the order of L-P-A (PcPAL4) or A-P-L (PsPAL1 & 2 of pepper, *Pisum sativum* and PvPAL2 of bean, *Phaseolus vulgaris*) (Table 4.2). The arrangement of the three boxes may have an influence on expression patterns of the PAL genes. In parsley, PAL1, PAL2 and PAL3, with the similar arrangement of the three boxes (P-A-L), were expressed similarly in various tissues; whereas PAL4 with a different arrangement (L-P-A) of the boxes showed different expression patterns (Logemann *et al.*, 1995). For example, parsley PAL4 was expressed at much higher level in roots and young stems than the other three PAL genes, which were expressed at much higher level than PAL4 in young leaves (Appert *et al.*, 1994). In parsley suspension cells, PAL4 did not respond to light, whereas the other PAL genes did (Logemann *et al.*, 1995).

A domain (AGGTGGCTGTATCGGTTGGAGA) located at -425 to –405 of the MePAL promoter was found to be of high similarity to the sequence of jasmonate acid (JA) responsive element (TGGAGGCTGGTTGGAGA) (Pastugila *et al.*, 1997). This element was identified in the promoter sequence of a wounding and bacterial infection-inducible receptor-like kinase gene SFR2 in *Brassica oleracea* (Pastugila *et al.*, 1997).

Two phytohormonal inducible elements, an auxin inducible, AATAAG, and an ethylene inducible element, GCCGCC (Eyal *et al.*, 1993; Liu *et al.*, 1994), were found at -2204 to –2199 and -2529 to –2524 respectively.

Inverted-repeated or palindromic sequences have been found to occur in both prokaryotic and eukaryotic genomes. Such repeated sequences are usually short and present at several functionally important regions in the genome. Site directed mutation in either or both of the direct repeat motifs in the promoter of a chalcone synthase gene in pea resulted in the reduction or loss in the ability for a nuclear binding factor to bind to a 61 bp region containing the repeats. This region was required for the maximal promoter activity and possibly for elicitor-mediated activation in pea (Seki *et al.*, 1996). In the MePAL2 promoter, there were also inverted repeats ACTCTTACAATCATTCTCA (-846 to -826), GCACCACCACCACG (-346 to -339), TCTCTTCACCTCTACTTCTCT (-177 to -164),

Table 4.2 Putative P, A, and L *cis*-elements and their positions on various PAL and 4CL gene promoters

Promoter	Box P	Position	Box A	Position	Box L	Position	Ref.
MePAL2*	TTTCAACCAACCCA	-248	CCCTCC	-219	GCTTACCTACC	-209	This thesis
PdPAL1	TTCTCACCAACCAC	-269	?		GCTTACCTACC	-89	Mitsumune 1999
PdPAL2	TTCTCACCAACCAC	-255	?		GCTTACCTACC		as above
Pc PAL1	CTCCAACAAACCCC	-175	CCGTCC	-160	TCTCACCTACC	-107	Logemann, 1995; Lois, 1989
Pc PAL2	CTCCAACAAACCCC	-175	CCGTCC	-160	TCTCACCTACC	-107	as above
PcPAL3	TTCCAACAAACCCC	-178	CCGTCC	-163	TCCCACCTACC	-98	as above
PcPAL4	CTCCAACCATCCAA	-183	CCGTTC	-136	TCTCACCAACC	-268	as above
Pc4CL1	CTTTACCAACCCCC	-52	CCGTCC	-228	TCTCACCAACC	-128	as above
AtPAL1	TCTCAACCAACTCC	-135	CCTCT	-260	GCTTACCTACC	-51	Ohl, 1990 & Warner,1995
AtPAL2	TCTCACCCACCCCT	-132	?		CCTTACCTAAC	-55	as above
LePAL5*	TTCCTACAACCCCC	-235	?		CTTTACCTACC	-207	Lee , 1992
PsPAL1	ATTCAACAAACCAC	-136	CCATCC	-265	CCTCACCTACC	-92	Yamada,1992 & 1994
PsPAL2	CTCAACCAAACCAC	-204	CGCTCC	-227	CTTTACCTACC	-151	as above
PvPAL2	TCTCCACCAACCCC	-123	ACGTCC	-274	ACCCACCTACC	-76	Cramer, 1989
PvPAL3	CAACCACCTACCCC	-77	?		ACTCACCAACC	-53	as above
StPAL1	TTCCAACAACCACC	-165	CCCTCT	-136	TCCCATCTCCA	-83	Joos,1992
StPAL2	CTTCAACAACCACC	-169	CCCTCC	-140	TCTCATCTACC	-72	as above
St4CL1	CTTTCACCTACCAC	-57	CTGTCC	-476	TCTCACCAACC	-170	as above
St4CL2	CTTTCACCTACCAC	-58	CTGTCC	-476	TCTCACCAACC	-168	as above

Note: Me, *Manihot esculenta* Crantz; Pd, *Populus trichocarpa* x *P.deltoides*; At, *Arabidopsis thaliana*; Le, *Lycopersicon esculentum*; Pc, *Petroselimum crispum*; Ps, *Pisum sativum*; Pv, *Phaseolus vulgaris*; St, *Solanum tuberosum* ? : not documented. * the positions of the boxes are numbered with MePAL or LePAL5 translation start site ATG as +1.

TTCCTCTCTCAGATCCTTTGTTTCTCCTT (-85 to -63) and direct repeats AATTAGC(AAATTCGGTT)AATTAGC (-753 to -730).

In addition to discrete *cis*-acting elements, some less well-defined domains in promoters also play important roles in gene expression. An AT-rich DNA-binding domain that occurs three times in mammalian high-mobility-group I/Y chromosomal proteins has recently also been identified in DNA-binding proteins from plants. It is suggested that the presence of the AT hook proteins may affect the expression of genes that determine the differentiation status of cells (Meijer *et al.*, 1996). In the MePAL2 promoter, the first 480 bp of the promoter was A/T –rich, with some regions consisting completely of A/T such as the region between –901 to –871 and –683 to –650. Interestingly, in the MePAL2 promoter the sequence immediately after the A/T-rich region becomes rich in either G/C (538-633, 56% GC) or C (738-769, 54.8%; 778-803, 56%).

4.3 Designation and construction of MePAL2 promoter-reporter gene constructs

Promoter-report gene fusion is widely used for determining the expression patterns of plant genes. The one of the most popular reporter genes in use is β-glucuronidase or GUS (*uidA*) gene first reported by Jefferson *et al.* (1987), which is a gene of bacterial origin whose expression can be driven in transgenic plants by a range of promoters. Endogenous GUS activity in plants is extremely low. Numerous plant genes have been successfully analysed using GUS as a reporter. Therefore, it was decided to characterize the MePAL2 promoter using promoter-GUS fusions.

The nopaline synthase (*nos*) terminator act as the site for polyadenylation and they are normally located within an expression cassette into which the promoter of interest can be inserted. One of the vectors used in this research, pKG, contained the *nos* terminator after the GUS gene. Whilst another vector, pMON977 used E9 3' as a terminator. E9 3' was identified by Hunt (1988) and used successfully in the research of cassava transformation system by Schopke *et al.* (1996) as a terminator (see Figure 2.1).

The Neomycin phosphotransferase (*nptII*) gene, which encodes resistance to aminoglycoside antibiotics such as kanamycin, neomycin, geneticin and paromomycin,

has been commonly used as a selectable marker in transgenic plants (Bevan *et al.*, 1983; Fraley *et al.*, 1983). This marker gene was used in plasmid pMON977 and in these experiments. This gene was driven by the cauliflower mosaic virus promoter (CaMV 35S) (see Figure 2.1), which is highly active in plant cells and directs constitutive gene expression in most tissues of transgenic plants.

As was shown in Figure 4.1, most putative *cis*-acting elements in the 5' flanking region of the MePAL2 were located –913 of the translation start site; therefore, promoter analysis using the GUS reporter gene was based on this region.

4.3.1 MePAL 840GUS/pUC and MePAL840GUS/MON

The 913 bp promoter in p1A26 (one subclone of MePAL2) was analysed for restriction enzyme sites that were suitable for recombination of the promoter into vectors for plant transformation. It was shown in Figure 3.7 (Chapter 3) that *Hin*dIII was at the beginning of the p1A26 insert while *Xho*II site was upstream and close to ATG. Therefore, they were suitable restriction sites to cut out the promoter for GUS fusion. However, there was no *Xho*II cloning site in the pKG vector, which contained GUS-*nos* terminator. To solve this problem, the *Xho*II site in the *Hin*dIII/*Xho*II promoter fragment needed to be blunt-ended or polished before being cloned to *Hin*dIII/*Sma*I sites of the GUS expression vector pKG to obtain PAL840GUS/pUC.

The recombination procedures of the construct PAL840GUS/pUC is shown in Figure 4.2. Firstly, clone p1A26 (1A26 fragment of MePAL2 in pUC19) was digested with the *Xho*II restriction enzyme, which released a 1800 bp fragment containing the *Xho*II/*Hind*III promoter region and about 900 bp of pUC19 vector. This fragment was gel-purified and then blunt-ended by mung bean nuclease. The blunted fragment was digested with *Hin*dIII to release the promoter from the sequence of vector. However, as these two fragments resulted from the *Hin*dIII digestion were very close in size, about 900 bp, it was difficult to separate them by gel electrophoresis. Therefore, these two bands were purified together and ligated to *Hin*dIII/*Sma*I cloning sites of pKG, which should result in clones either containing the 840-PAL promoter or the 900 bp pUC19. After transformation, plasmid DNA from six colonies were digested with

*Hin*dIII/*Eco*RI and tested by Southern hybridisation to identify the ones containing MePAL 840-GUS construct. Southern analysis showed two of these clones hybridised to MePAL2 probe covering the promoter region (Figure 4.3). One of the positive clones was sequenced, which confirmed the expected MePAL 840-GUS sequence. This construct containing MePAL 840-GUS promoter, was named MePAL840GUS/pUC.

One of the vectors successfully used in cassava transformation is pMON977 (Gonzalez, 1998), which has E9 3' terminator and *npt*II gene, but not the GUS gene. The MePAL840-GUS cassette was cloned into a modified pMON977 (Figure 4.4). In detail, PAL840GUS/pUC was initially digested with *Hin*dIII/*Sst*I and the two fragments released, 840GUS and pUC19 vector plus *nos* terminator, were all about 2700 bp. It was difficult to separate these two fragments by gel electrophoresis. Therefore, these two fragments were purified together and then inserted to pMON977 after the cassava mosaic virus (CVMV) promoter in the vector was removed by *Hin*dIII/*Sac*I digestion. The recombinants were grown on LB plates supplied with streptomycin which is the bacterial selectable marker of pMON977 (for the map of pMON977 see figure 2.1). These recombinants should contain either MePAL840/GUS or pUC19+*nos* terminator. In order to identify the clone containing MePAL840/GUS, these recombinants were tested on LB medium plates supplied with ampicillin. If the clones contained the pUC19, they would grow as there is an ampicillin resistant gene in pUC19. Therefore, those clones that did not grow on the ampicillin medium should contain the PAL840GUS fragment. One of these clones that did not grow on ampicillin was sequenced and confirmed to be PAL840/GUS in pMON977, designated PAL840GUS/MON (Figure.4.4).

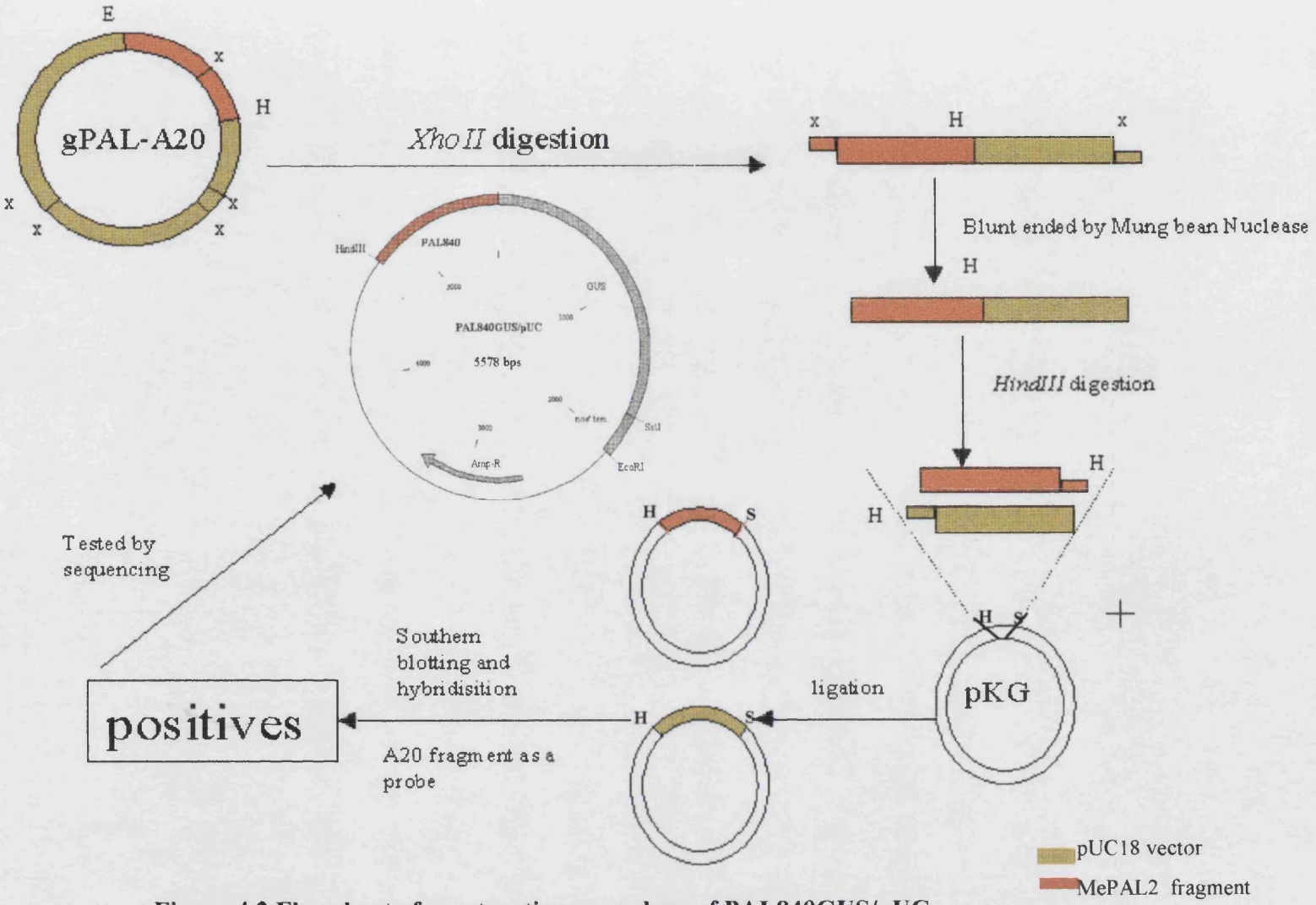


Figure 4.2 Flowchart of construction procedure of PAL840GUS/pUC.
X: *Xho*II site; **H:** *Hin*dIII site; **E:** *Eco*RI site; **S:** *Sma*I site

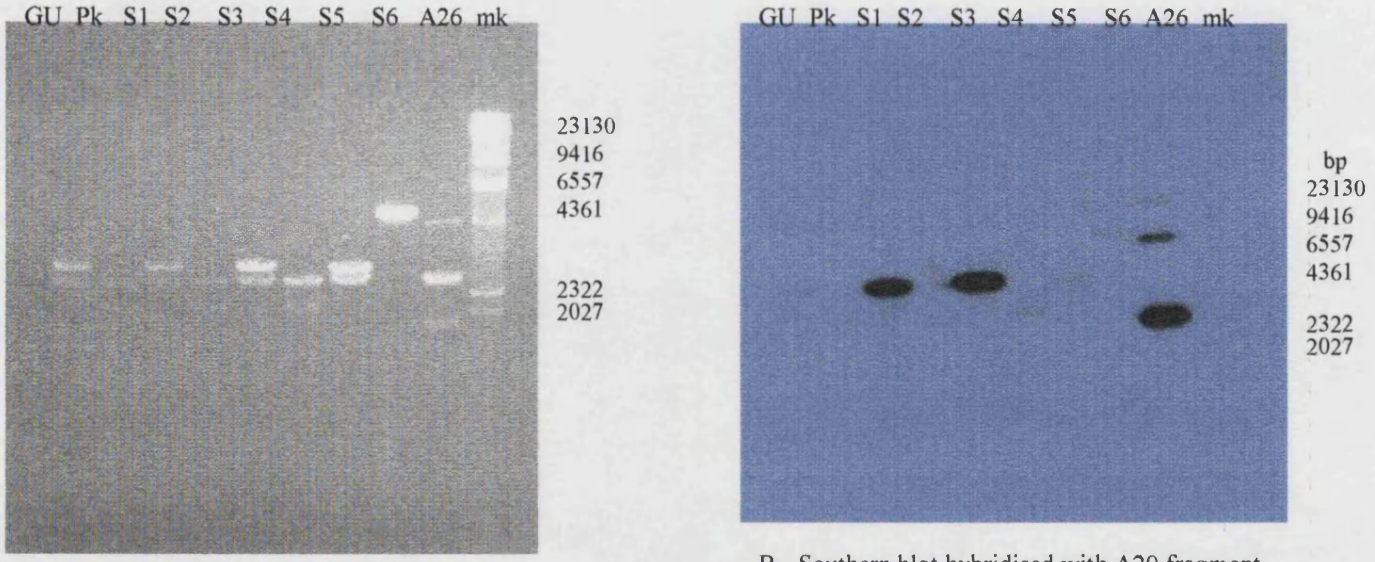
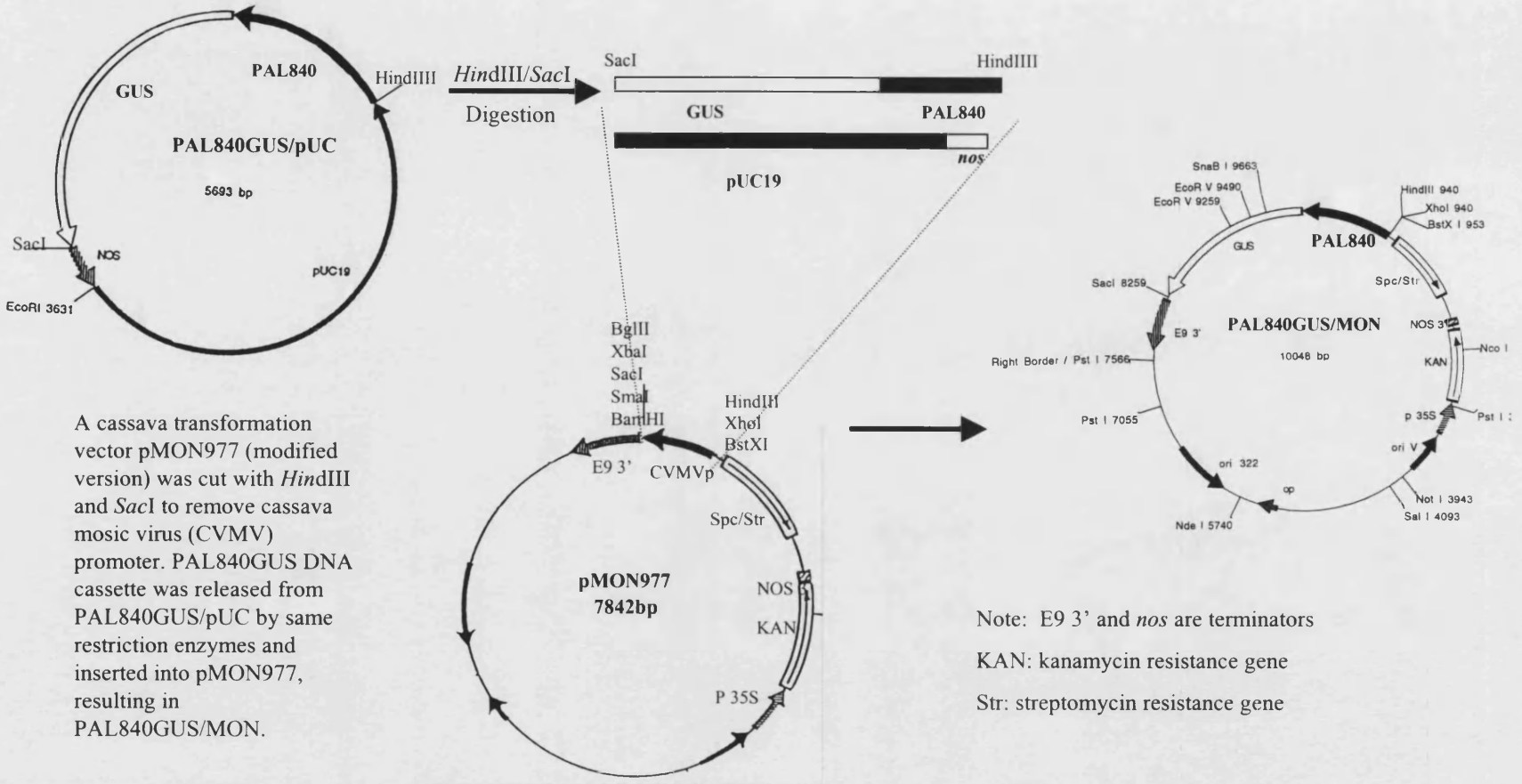


Figure 4.3 Selection of positive clones of MePAL840GUS/pUC by Southern hybridisation.

Plasmid DNA from six clones (S1 to S6), resulted from ligation of two *Hin*dIII/*Xho*II (blunt-ended) fragments into pKG, were run in agarose gel (A), Southern-blotted and hybridised with a probe (A20) made with *Hin*dIII / *Eco*RI fragment of pA26 (B). GU-CaMV 35S-GUS plasmid as size control, pK-pKG as a negative control, A26-pA26 plasmid digested with *Hin*dIII / *Eco*RI , mk-*Hin*dIII cut λ DNA as a size marker.



A cassava transformation vector pMON977 (modified version) was cut with *Hin*dIII and *Sac*I to remove cassava mosic virus (CVMV) promoter. PAL840GUS DNA cassette was released from PAL840GUS/pUC by same restriction enzymes and inserted into pMON977, resulting in PAL840GUS/MON.

Note: E9 3' and *nos* are terminators

KAN: kanamycin resistance gene

Str: streptomycin resistance gene

Figure 4.4 Construction flowchart of PAL840GUS/MON.

4.3.2 Truncated MePAL2 promoter constructs

Analysis of promoter activity in transgenic plants using GUS as a reporter gene often centres on the construction of transgenes with various regions of the promoter either deleted or mutated. The reasoning behind such an approach is to determine which regions of the promoter DNA function in regulating the transcription of the gene. This is a tried and tested method and has been of great benefit in defining control regions in a large number of plant promoters. For example, P, A, and L boxes were identified as *cis*-acting elements in phenylpropanoid genes such as PAL and 4CL in many plants (Lois *et al.*, 1989; Logemann *et al.*, 1995). These three *cis*-elements were also identified in the MePAL2 promoter (Table 4.2). To determine the functional properties of these *cis*-elements in cassava, the following MePAL2 promoter deletions were carried out in ILTAB, Danforth Plant centre, USA.

PCR amplification has been one of the most used techniques for promoter deletion and mutation. To facilitate the PCR cloning, restriction sites for *Bam*HI and *Hind*III were incorporated into the primer by attaching the related sequence to the primer or mutating some of the bases in the primer (*Bam*HI and *Hind*III sites are underlined and the mutated base pairs were in bold as shown in the primer sequences described below) in all of primers used for constructing the deletions. The first MePAL2 promoter deletion construct was designed to contain all P, A, L boxes and CAAT box but not the G box. Primers, PD1 (-424 to –398) GGTGGCTGA**A**AGC**T**TTGGAGACCCAATC and PDR (reverse primer, -57 to –89) GAGAAACAAAGGATC**C**GAGAGAGGAAATC were used to amplify that region from p1A26 plasmid containing MePAL2 promoter, and resulted in the PCR product, PAL400. PAL400 DNA was purified using gel-band purification, *Bam*HI and *Hind*III digested and cloned into a modified pMON977 (this version containing the GUS gene in the vector), yielded PAL400GUS/MON (Figure 4.5). Attempts were also made to clone PAL400 into pKG for transient assays, but these were not successful.

Similarly, primers PD2 (-247 to –226) TCAACCAACC**A**AGCTTCTCTCC and PDR were used to amplify the fragment that containing the A and L boxes but not the P box.

This PCR product, PAL260, was purified and cloned into the pKG vector, leading to construct PAL260GUS/pUC (Figure 4.5).

Another construct PAL200GUS/pUC was obtained by PCR using primers PD3 (-219 to –192) CCTCCTTAAGCTTACCTACCCATTACAC and PDR. The PCR product, PAL200, containing only the L box and the two putative TATA boxes, was cloned into the pKG vector (Figure 4.5).

Before the construction of PAL200GUS/pUC and PAL260GUS/pUC, attempts at cloning PCR products PAL260 and PAL200 into pMON 977 were made, but without success.

4.3.3 TATA box mutation constructs

TATA boxes are found in all prokaryotic and eukaryotic gene promoters, they are involved in the orientation of RNA polymerase II. It has been shown that mutations of the TATA box can lead to a substantial reduction of promoter activity (Watson *et al.*, 1992). Two TATA box like motifs were found in the MePAL2 promoter. In order to identify which putative TATA box is the functional TATA box, two constructs TAM1-GUS and TAM2-GUS were made to mutate the first or the second TATA box, from TATA or TATTTAAA mutated to TCTC or TCTTTCCC (changed nucleotide A to C) respectively. To facilitate the PCR cloning, restriction sites for *Bam*HI and *Hind*III (underlined in the primer sequence described below) were incorporated into the primers (in bold in the primers described below) in some of primers used for constructing the TATA boxes mutation. The plasmid pA26, containing MePAL2 promoter, was used as template.

The construction of TAM1 involved three PCR amplifications (Figure 4.6). Firstly, primer PTACF (CACTCGAGTGGAAGCTTATTGTG, -925 to –903) covering part of the 5' end of the promoter and of pUC vector, and PTR1 (GAAGAGACTT CC**AGAG**AAGTGTGTAATGGGT, -172 to –202) covering the first TATA box but with the TATA mutated to TCTC (complementary bases of mutated box underlined and mutated bases in bold), were used to produce a PCR fragment TM1. Similarly, a

forward primer PTF1 ACCCAT TACACACTTCTCTGGAAGTCTCTTC (–202 to-172) (mutated TATA box underlined and mutated bases in bold), which contained the mutated TATA box, and a reverse primer PDR (-57 to –89) GAGAAACAAAGGATCCGAGAGA GGAAATC, were used to amplifying the second PCR fragment TM2. TM1 covered the region between –913 to –172 of the promoter and contained the first TATA (mutated to TCTC), whereas TM2 covered the region between –172 to –57of the promoter and contained the same TATA box mutation with the second TATA box unchanged.

Since these two PCR fragments TM1 and TM2, overlap at the first TATA box, they can be integrated into one fragment by PCR or by over-lap PCR extension. So the 840 MePAL2 promoter with the first TATA box mutated to TCTC was amplified with PTACF and PDR primers from the mixture of purified TM1 and TM2 DNA as templates (at the same molecular ratio). This PCR fragment was designated TAM1. The TAM1 fragment was then purified and digested with *Hin*dIII and *Bam*HI. The purified TAM1 with *Hin*dIII/*Bam*HI cloning sites were then inserted into the same sites of PKG to yield TAM1-GUS (Figure 4.6 and Figure 4.7). The positive TAM1-GUS clone was confirmed by sequencing.

TAM2-GUS containing the mutated second TATA box (TATTTAAA to TCTTTCCC on -143 to –136 of 5' MePAL2 promoter) was constructed in the same way as the construction of TAM1-GUS. The primers used were PTACF/PTR2 (PTR2: GAAGGAGTGGA**GGGG**AAA**G**AGAAGGTGGGGAG –126 to –155) and PTF2/PDR (PTF2: CTCCCCACCTTCT**C**TTT**CCCC**TCCACTCCTTC –155 to –126) to amplifying two mutated second TATA box fragments. Similar to TAM1, primers PTACF and PDR were used for amplifying the second TATA box mutation of 840 MePAL2 promoter. The structure of the promoters with mutated TATA boxes is shown in Figure 4.6. The positive TAM2-GUS was also confirmed by sequencing.

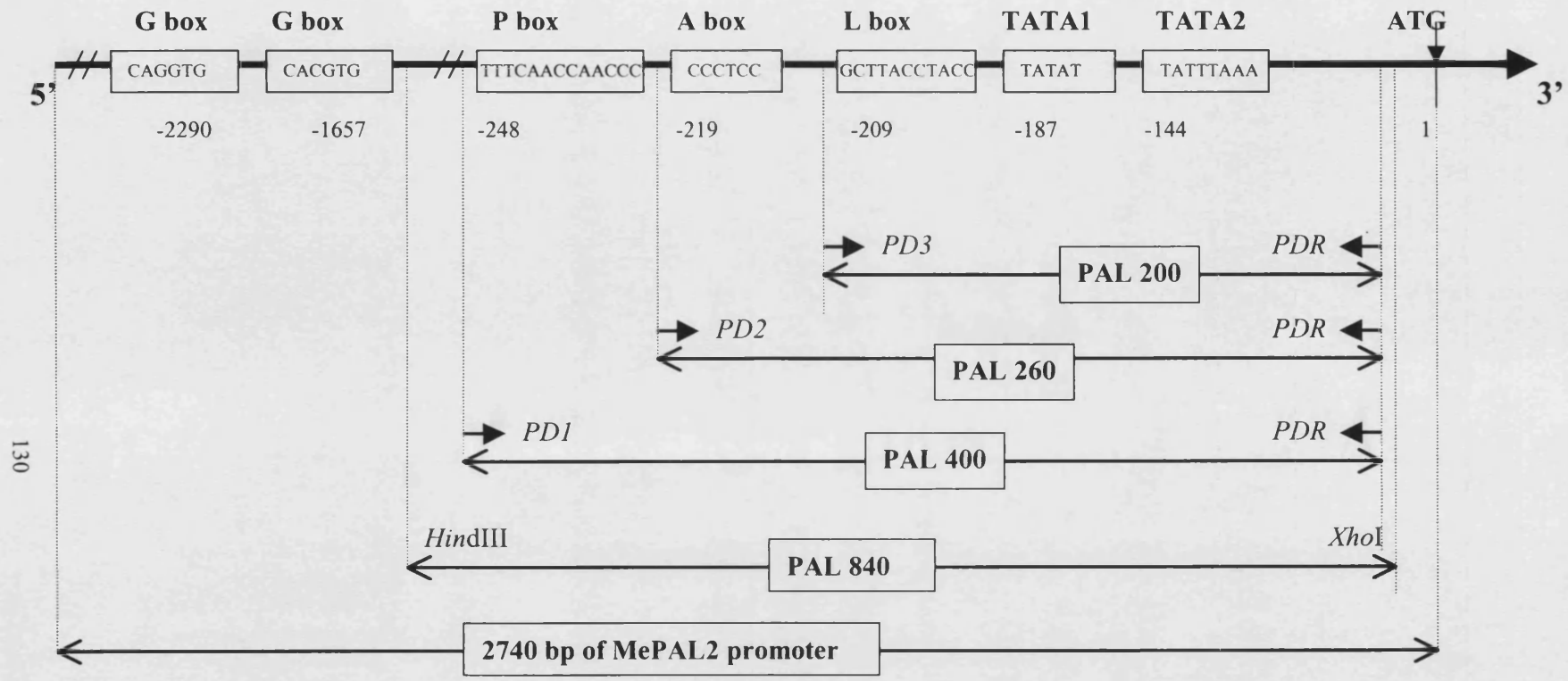


Figure 4.5 Linear map of PAL promoter truncations.

Three promoter truncations, PAL 200, 260 and 400 were made using PCR to amplify the corresponding regions of MePAL2 promoter. PAL800 fragments was made from *Hin*dIII / *Xho*I digestion of 1A26 MePAL2 clone. The truncated promoters then were cloned into pMON977 or pKG to drive GUS expression. G, P, A & L boxes are putative *cis*-elements, which are conserved in PAL and 4CL promoters. TATA1 & TATA2 are two putative TATA boxes. PD1, PD2, PD3 and PDR are PCR primers with restriction sites. ATG is putative translation start site. Positions of sequence motifs are indicated by nucleotide numbers upstream of putative translation start site (ATG). Primers are in italic.

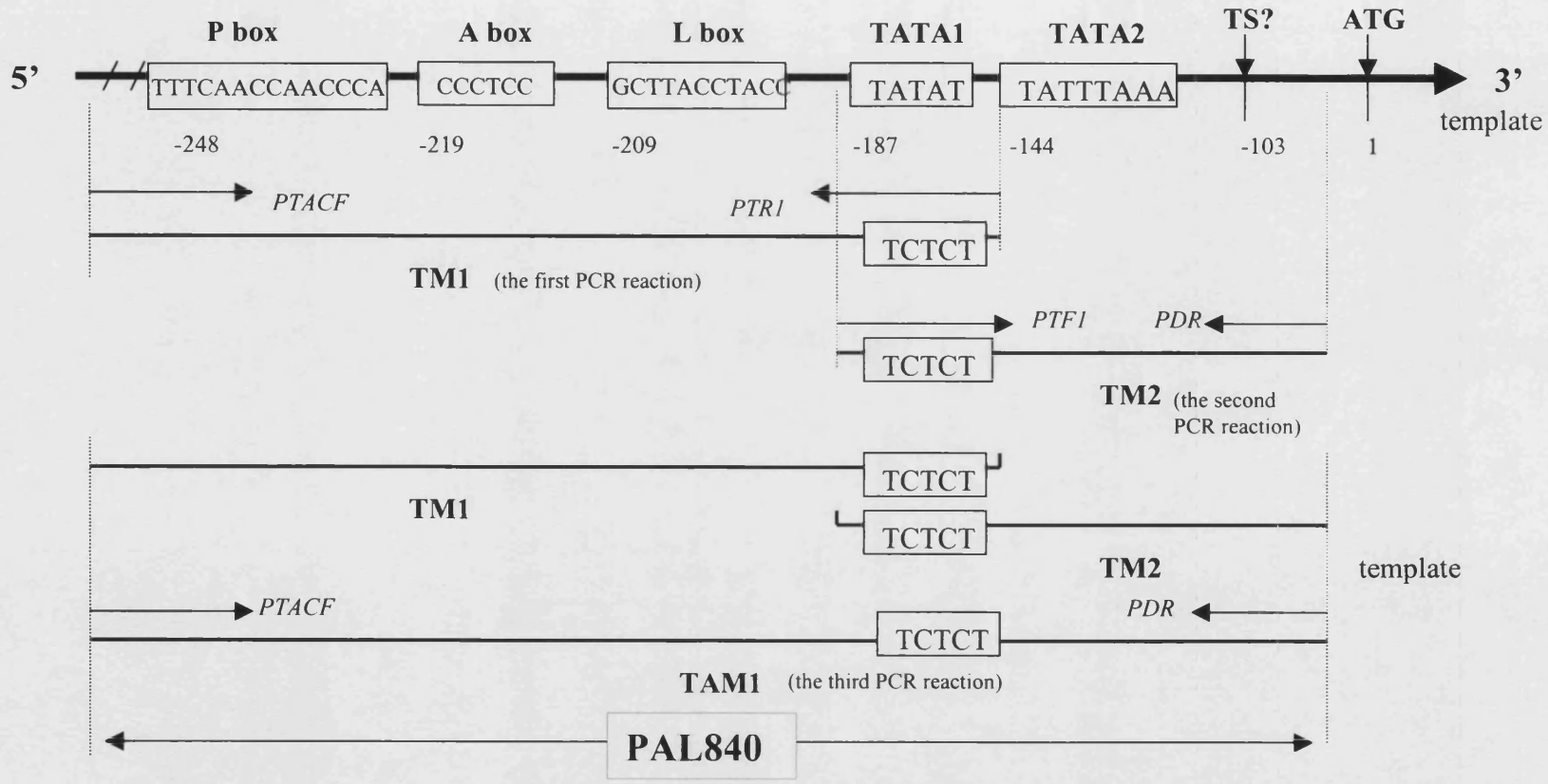


Fig 4.6 Construction of TAM1, a MePAL2 promoter with mutated TATA box.

The mutation was made by two rounds of PCR reactions. The first round PCR consisted of two sets of PCR reactions amplifying TM1 and TM2 respectively. TM1 was amplified using *PTACF* and *PTR1* (with TATAT mutated to TCTCT). TM2 was amplified using *PTF1* (complementary to *PTR1*) and *PDR*. The second round PCR was carried out with *PTACF* and *PDR* using the mixture of TM1 and TM2 as template, resulting in the PAL840 promoter with TATA1 mutated to TCTCT.

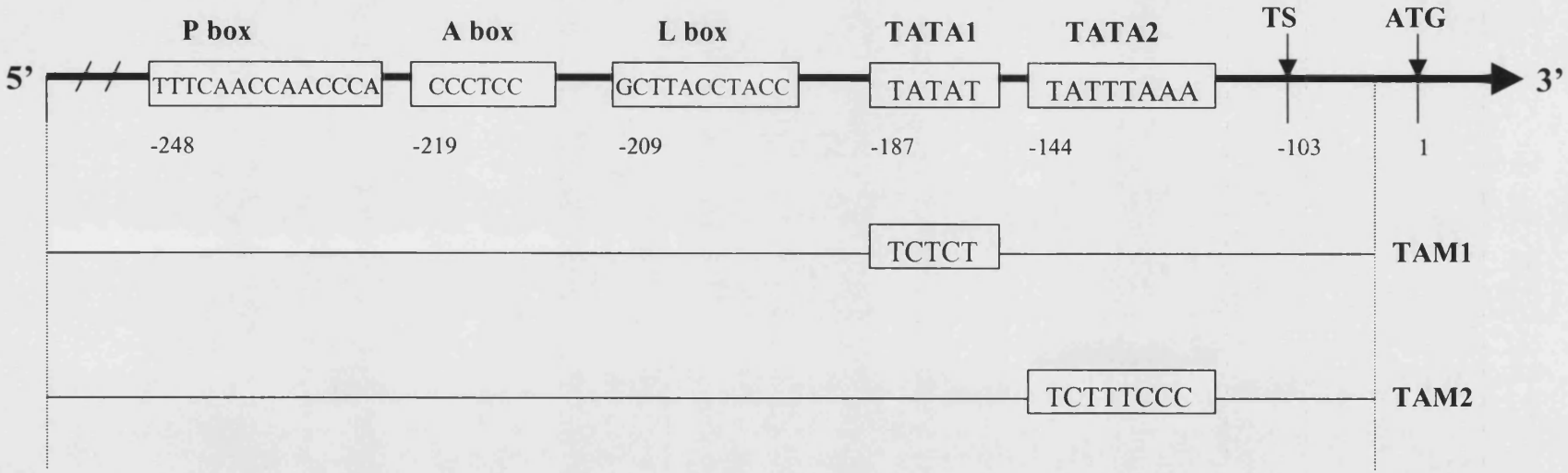


Figure 4.7 Summary of TATA box mutations of the MePAL2 promoter.

The TATA1 mutation was made as described in the legend to Figure 4.6. The mutation of the second putative TATA box was carried out in the same way as the TATA1 mutation, in which TATTTAAA was mutated into TCTTTCCC.

4.4 Transient activities of the MePAL2 promoters in cassava

Transient assays have been widely used to study gene regulation in animal cells and in protoplasts from monocots and dicots (Werr and Lorz, 1986; Lipphardt *et al.*, 1988; Jones *et al.*, 1989). They are particularly valuable tools for analysing gene expression in plants for which efficient transformation procedures do not yet exist. Transient assays using marker genes (such as the GUS gene) by particle bombardment is a fast, easy and efficient approach to obtain initial information about the activity of a promoter. In cassava, folded leaves and friable embryogenic callus (FEC) were commonly used target tissues in transient assays (Schopke *et al.*, 1996).

In order to test the activities of putative *cis*-acting elements in the MePAL2 promoter, different truncated MePAL2 promoter constructs were transformed into cassava embryogenic suspension cells and / or folded cassava leaves by particle bombardment and transient GUS activities were analysed. These constructs include PAL840GUS/pUC, truncated promoter constructs PAL400GUS/MON (it would have been better to use a PAL400GUS/pUC construct for comparison but attempts to construct it failed), PAL260GUS/pUC and PAL200GUS/pUC, and the constructs with mutated TATA boxes, TAM1-GUS and TAM2-GUS. The cauliflower mosaic virus promoter (CaMV 35S) promoter has been widely used in plant biotechnology as this promoter is highly active in plant cells and directs constitutive gene expression in most tissues of transgenic plants (Schopke *et al.*, 1996; Sanders *et al.*, 1987). In this research, CaMV 35S promoter-GUS p35S (in pU19) was used as a positive control while the promoterless GUS-terminator plasmid pKG was used as a negative control.

4.4.1 Transient activities of PAL 840 in folded cassava leaves

Young leaves, especially folded leaves, are ideal target tissues for particle bombardment as they are a convenient target and are penetrated easily by DNA coated particles. Therefore, the transient activity of PAL840 promoter was tested in folded leaves (the technique of particle bombardment of cassava young leaves and embryogenic suspension cells was kindly demonstrated by Luong, H. T. in Dr Paul Lazzeri's laboratory in Rothamsted, UK) as a preliminary test of the construct and

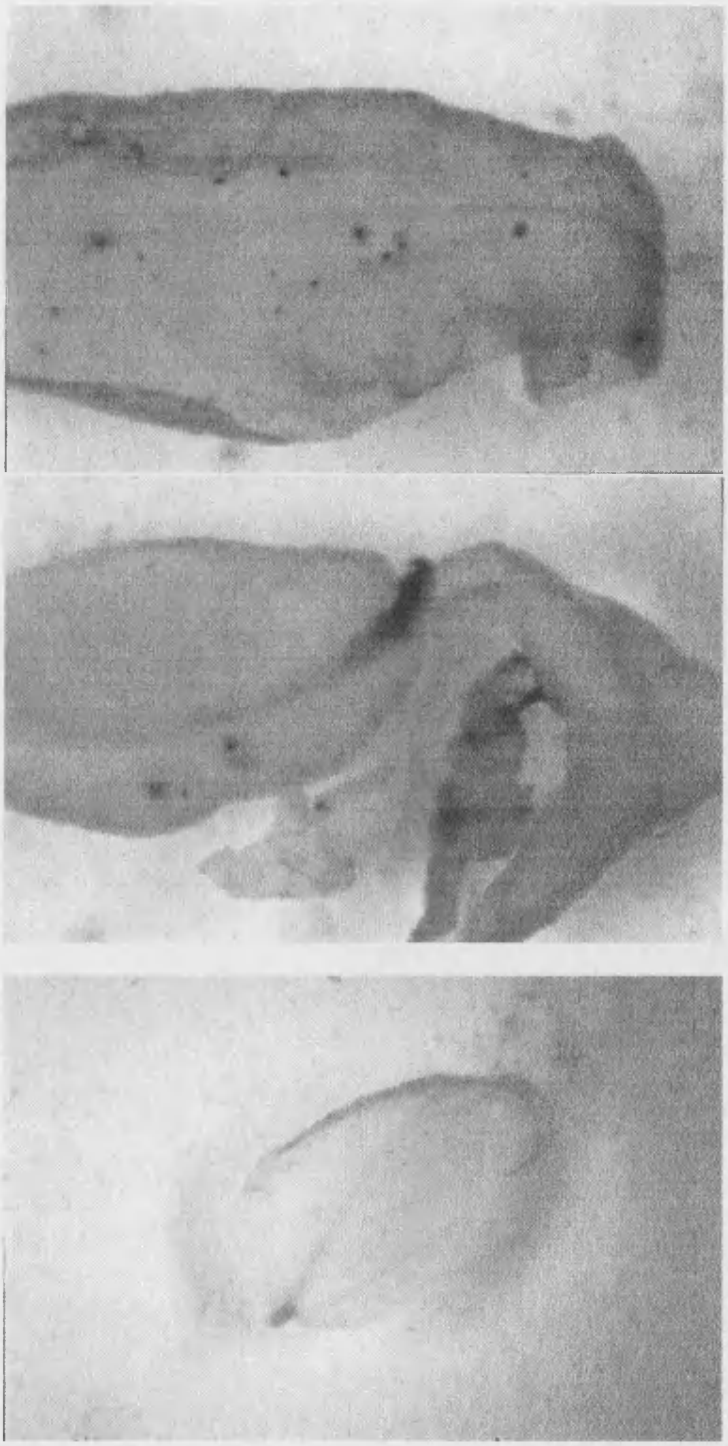
promoter before further work was done on the transient activities of the promoter in suspension cells and stable transformation in ILTAB, USA. PAL840GUS/pUC and two controls, pKG and p35S, were bombarded into folded young cassava leaves harvested from *in vitro* plants. For each construct, about 5 μg of DNA was bombarded into around 15 folded leaves. After bombardment the leaves were left in MS medium and incubated for two days at 27°C dark before staining. Then the leaves were stained in GUS staining buffer for 2 hours in the dark at 37°C, and washed with 70% ethanol (as described in chapter 2). From Figure 4.8, it is clear that GUS expression was much higher from p35S than from PAL840GUS/pUC. There were more than 40 blue spots in a folded leaf bombarded with p35S but only one blue spot was visible in a leaf bombarded with PAL840GUS/pUC. There were no blue spots visible in the negative control leaf bombarded with the pKG construct. However, it should be pointed out that the data obtained from the positive control as shown above suggests a poor delivery efficiency as 35S-GUS should give 10s to 100s blue spots in such an assay in cassava (Lazzeri, P., personal communication).

Even though immature leaves are a convenient target tissue for transient assays, they were not sufficiently homogeneous to provide reliable quantitative data. Therefore, the particle bombardment of cassava leaves were not further optimised for high delivery efficiency, instead, transient activities of the promoters were further analysed using suspension cell culture.

4.4.2 Comparison of the transient activities among truncated MePAL2 promoters in cassava suspension cells

Embryogenic suspension cells or new embryogenic units in cassava friable embryogenic callus were developed from the surface cells of the globular embryo clusters and seem to be of single cell origin, which makes them an ideal target tissue for transformation. Suspension cells are especially suitable for transient assay in that identical materials can be obtained easily for different treatments to minimise the experimental error. For each construct and control plasmid, about 5 μg DNA was bombarded into embryogenic suspension cells. The bombarded cells were incubated for two days (photo period 16 hours) at 27°C. The cells were stained in GUS staining

buffer for 2 hours in the dark at 37°C, and washed with 70% ethanol. Figure 4.9 shows the transient expression of GUS in suspension cells with PAL840GUS/pUC, pKG and p35S-GUS. As can be seen transient GUS expression of PAL840GUS/pUC in embryogenic suspension cells is much higher than in immature cassava leaves (Figure 4.8). The significant difference in the transient expression between suspension cells and leaves may lie with the nature of them, with embryogenic suspension cells more active than the cells in leaves.



A: Positive control
p35S plasmid

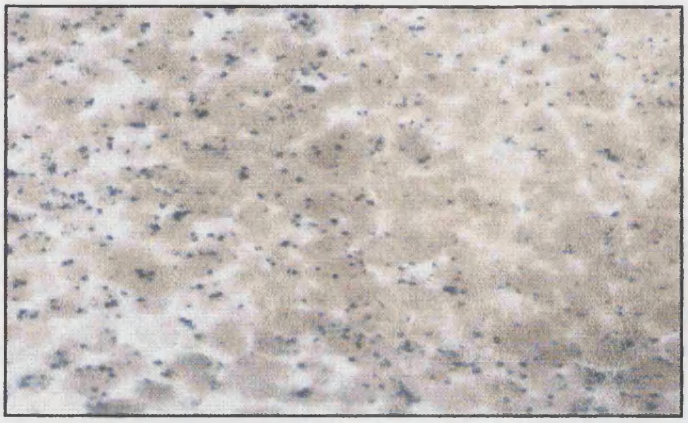
B: Test

PAL840GUS/pUC plasmid

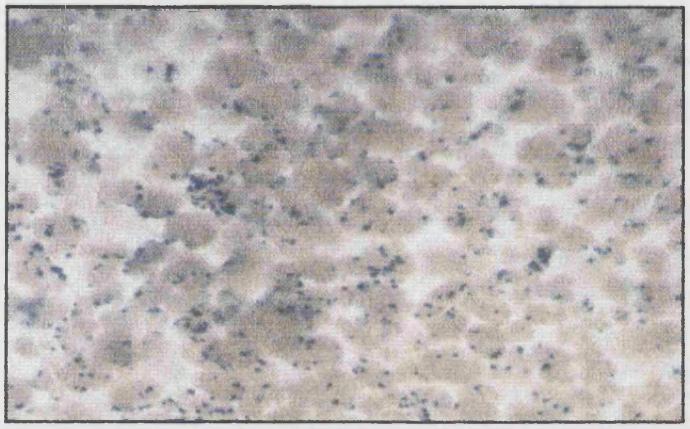
C: Negative control
Promoterless GUS gene plasmid--pKG

Figure 4.8 Transient GUS assay of MePAL2 promoter in folded cassava leaves.

PAL840GUS/pUC plasmid was bombarded onto *in vitro* folded cassava leaves and stained for GUS activity after incubated at 37°C in the dark for 24 hrs (B). Plasmid p35S (GUS driven by CaMV 35S promoter) and promoterless GUS gene plasmid pKG were used as positive (A) and negative (C) control respectively.



Positive control: p35S plasmids



Test:

PAL 840-GUS/pUC plasmids

Negative control: Promoterless GUS gene plasmids--pKG

Figure 4.9 Transient GUS assay PAL840 in cassava suspension cells.

PAL840GUS/pUC plasmid was bombarded onto cassava suspension cells and stained for GUS activity after incubation at 37°C in the dark for 2 hours (middle panel). Plasmid p35S (GUS driven by CaMV 35S promoter) and promoterless GUS gene plasmid pKG were used as positive (top panel) and negative (bottom panel) controls respectively.

The transient activities of truncated promoters were measured by the number of blue spots resulting from GUS expression. Table 4.3 shows the results of the GUS assays of the PAL 840 and truncated promoter constructs. It should be pointed out that the number of molecules of the plasmid PAL840GUS/pUC, PAL 260GUS/pUC, 200GUS/pUC, p35S and pKG bombarded was nearly twice as many as that of PAL400GUS/MON, since equal amount of plasmids were used but the size of the former was about half of that of the latter. The GUS assay showed that GUS expression in cells transformed with PAL840GUS/pUC is more than three or four times of that in cells transformed with the truncated MePAL2 promoter constructs PAL260 and PAL200. The GUS expression level driven by PAL840 promoter in the transient assay into suspension cell was high, which was as high as that driven by 35S promoter. The negative control pKG, which contained GUS without any promoter, showed no GUS activity. Therefore the *cis*-elements in each deleted constructs played important roles for reporter gene expression level.

Promoter	A	B	C	D	Mean
35S	3956	4218	3560	4828	4140±382
PAL840	4144	3660	4930	4728	4366±463
PAL260	1512	1026	927	-	1155±238
PAL200	1120	684	814	-	872±164
pKG	6	0	3	5	3.5±1.5
PAL400*	640	828	836	584	722±110

Table 4.3 Transient GUS assay (measured by the number of blue spots) of different constructs of truncated promters and control plasmids in suspension cells (5μg of plasmids for each reaction).

[a] A-D are from four separate cassava transformation events

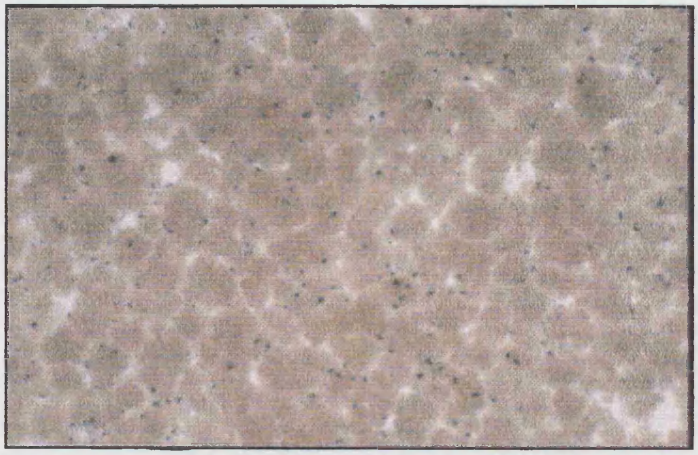
* The PAL400 plasmid here is much larger than the other constructs in size.

4.4.3 Identification of probable TATA box by transient activities of MePAL2 promoter constructs with mutated TATA boxes

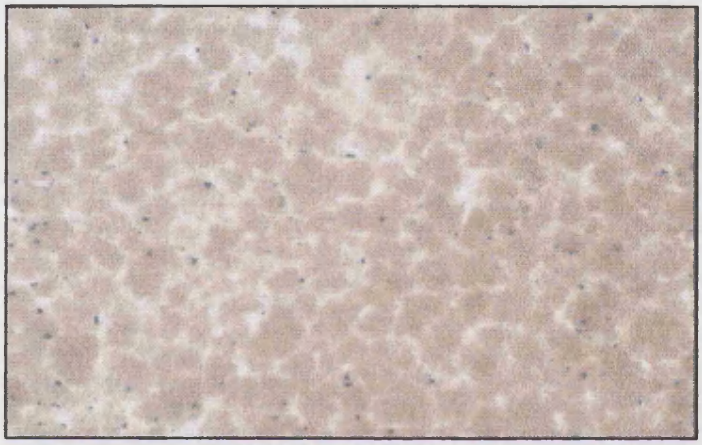
The two TATA mutation constructs TAM1-GUS and TAM2-GUS were bombarded onto suspension cells under the same conditions as described for the transient assay of truncated MePAL promoters (4.4.2). Both TAM1-GUS and TAM2-GUS showed transient GUS expression in suspension cells (Table 4.4), but TAM2-GUS showed about half of the expression of TAM1-GUS. The difference between TAM1-GUS and TAM2-GUS in GUS expression or density of the blue spots was also visible (Fig 4.10). This means that mutation of the second putative TATA box reduces significantly the transient activity of the promoter, which suggests that TATA2 is the actual TATA box.

Promoter	A	B	C	D	Mean
TAM1	2495	3053	3752	4054	3338±565
TAM2	1207	855	2040	2553	1663±633

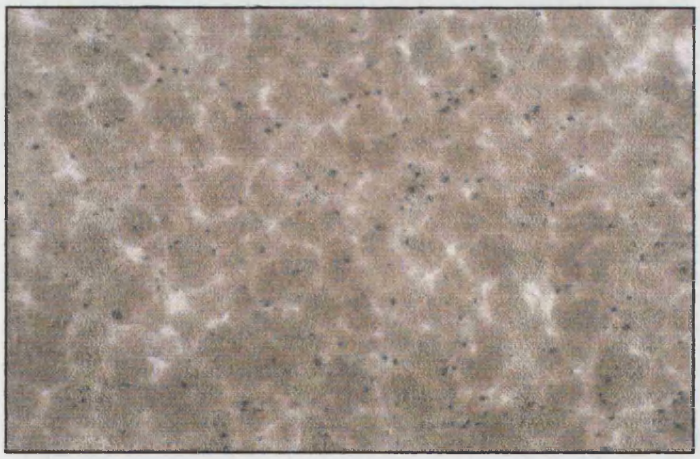
Table 4.4 Transient activities of two promoter constructs with mutated TATA boxes. Note: The activities of the promoter were measured by number of blue GUS staining spots derived from each bombardment experiment, in which 5 μg of plasmids was used. A, B, C and D were four individual experiments.



A: TAM1 mutation
Transient GUS expression of TATA1 mutated promoter in embryogenic suspension cells



B: TAM2 mutation
Transient GUS expression of TATA2 mutated promoter in embryogenic suspension cells



C: PAL840GUS/pUC
Transient GUS expression of PAL840GUS/pUC in embryogenic suspension cells as a control

Figure 4.10 GUS assay of the promoter constructs with mutated TATA boxes. TATA1- and TATA2-mutated promoter-GUS plasmids TAM1 (panel A) and TAM2 (panel B) were bombarded onto embryogenic suspension cells and stained for GUS activity after incubation at 37°C in the dark for 2 hrs. PAL840GUS/pUC plasmid used as a control. Promoterless and p35S plasmids controls see Figure 4.9.

4.5 Discussion

The analysis of the MePAL2 promoter has highlighted the presence of potential regulatory motifs throughout the 5'-flanking sequence of MePAL2, most of which are located within 1000 bp upstream of the translation start site. The presence of these motifs within the promoters of other genes have been shown to confer responsiveness on the promoter and, in some instances been shown to bind nuclear protein factors that regulated transcription. In order to understand the functional significance of some motifs and regions in the MePAL2 promoter, promoter deletion was carried out. GUS expression driven by 913 bp of the MePAL2 promoter and truncated promoters was analysed by transient assays.

Transient assays have been widely used to study promoters in plants (Jones *et al.*, 1989). Early successful examples of transient assays of promoters include UV-induction of gene expression from chalcone synthase promoter in parsley protoplasts (Dangl *et al.*, 1987), the ABA-induction of a wheat embryo gene promoter (Marcotte *et al.*, 1988) and gibberellin-induction of a wheat α-amylase promoter (Huttly and Baulcombe, 1989) in protoplasts.
In the present study, MePAL2 promoter fragments, including PAL840, PAL400, PAL260 and PAL200 were fused with GUS gene and transferred into embryogenic cassava suspension cells to test the transient activities of these truncated promoters. The truncated promoter PAL400, which excluded a AT-rich domain, an inverted and a direct repeat, box V, CCAAT box, half of the JA responsive element (Figure 4.11) and was about 400 bp shorter than PAL840, was still active. The truncated promoters PAL260 and PAL200, which excluded the region containing P box or P & A boxes showed much lower GUS expression than PAL840, indicating that the deleted regions are crucial for the high activity of the PAL840 promoter.

A putative TATA box, nearest the initiator ATG, is probably the genuine TATA box as indicated by transient expression assays of promoters mutated at putative TATA boxes.

The transient activity, however, cannot completely represent the properties of the promoter. Significant differences between transient and stable activities of promoters

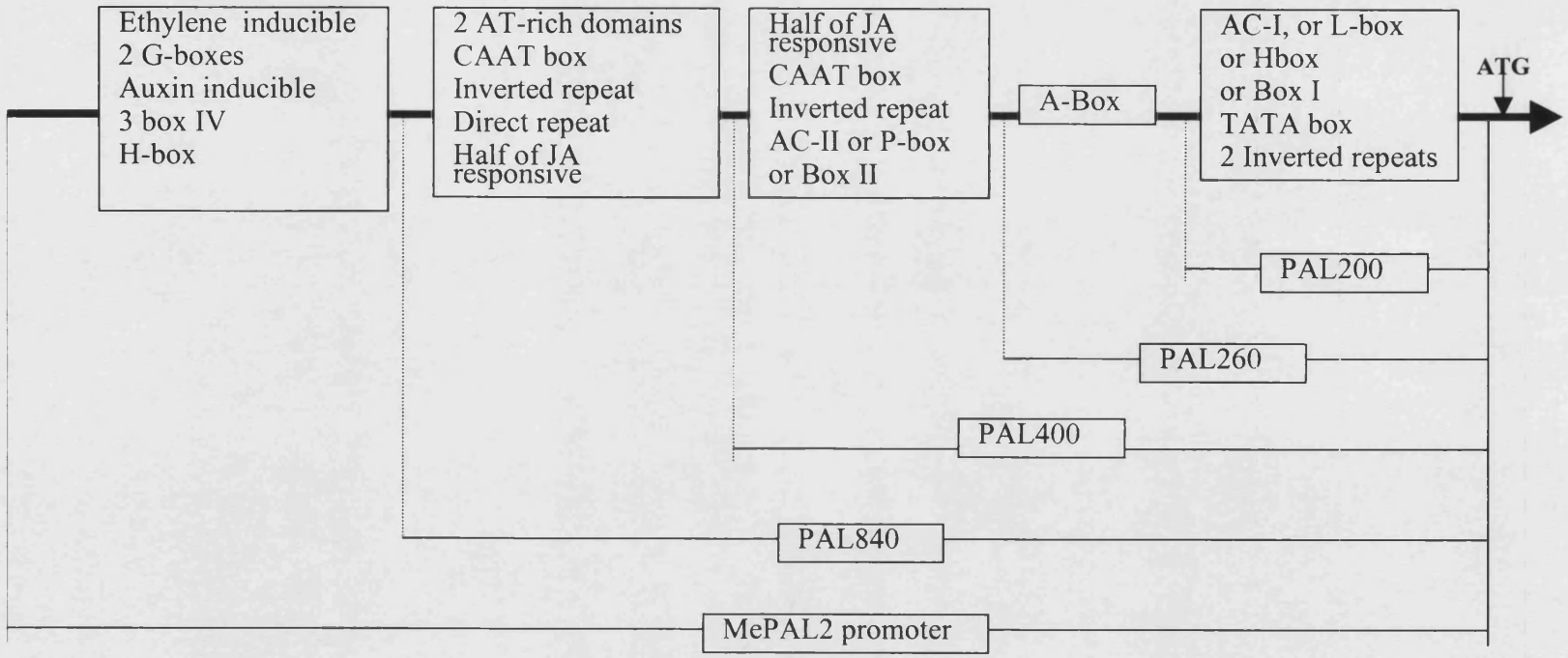


Figure 4.11 *cis*-elements and motifs in different regions of the MePAL2 promoter.

cis-elements and motifs identified in the MePAL2 promoter are presented in relation to the truncated constructs.

have been observed. For example, the region between –335 to +28 of the promoter of pathogenesis-related (PR) protein PR-1a gene from tobacco was highly active in transient assays, but did not give any detectable expression of the GUS reporter gene in transgenic plants under various conditions applied (Beilmann *et al.*, 1991). The region between –343 and –208 of the CaMV 35 S promoter had no apparent effect on the level of luciferase expressed in transient assays in carrot protoplasts (Ow *et al.*, 1987) but the same region was found to be responsible for ~40 to 50% of the total activity in leaf tissues in transgenic tabacco plants (Fang *et al*, 1989). The putative *cis*-element P-A-L boxes, essential for the activity of PAL promoter (Hatton *et al.*, 1995), were partially deleted in PAL260 (without P) and PAL200 (without P-A) but these truncated promoter showed transient, though reduced, activity. Stable transformation of cassava by these MePAL2 promoter constructs would help to clarify the importance of these elements in the activity of MePAL2, more importantly, would reveal the temporal and spatial expression patterns of the gene.

CHAPTER FIVE: GUS EXPRESSION PATTERNS DIRECTED BY THE MEPAL2 PROMOTER IN TRANSGENIC CASSAVA AND RICE

5.1 Introduction

Genetic transformation is a powerful tool for understanding complex biochemical pathways and analysing gene regulation by expressing reporter genes driven by a promoter. With the establishment of a cassava transformation system by particle bombardment, cassava transformation is being used to explore the possibility of genetic modification of cassava and to test the regulation of gene promoters. For example, transgenic cassava plants containing CsCMV coat protein gene and ACMV AC1 replicase gene were produced to obtain lines with resistance to CsCMV and ACMV (Taylor, 1999). A few lines showed resistance to CsCMV in a greenhouse (Taylor N, personal communication). CaMV 35S and a cassava vein mosaic virus (CsVMV) promoter driving the GUS gene have been transferred into cassava for optimizing the genetic transformation system and studying the regulation of the CsVMV promoter (Verdaguer *et al.*, 1998).

In the last chapter, a number of interesting motifs such as putative *cis*-acting elements were identified in the sequence of the MePAL2 gene promoter by comparing with known motifs from genes in other plant systems. Based on these analyses, a MePAL2 promoter deletion series and TATA boxes mutation were made and fused with the GUS gene. Transient assays of different constructs showed that truncation of the promoter reduced GUS activity, and the constructs PAL260 and PAL200 were transiently functional, though P-A-L box was partially deleted in these promoter constructs. These elements however, were shown to be essential for the expression of bean PAL promoter in tobacco (Hatton *et al.*, 1995).

In this chapter, stable transformations of cassava with the promoter deletion-GUS series were carried out and transgenic cassava lines were obtained and analysed for the temporal and spatial activities of the MePAL2 promoter during plant development and the wounding response in the cassava storage roots.

It has been shown that monocot PAL promoters (such as rice) can drive reporter genes in dicots such as tobacco (Zhu, *et al.*, 1995). However, there is no report on whether a dicot PAL promoter is able to drive reporter gene expression in a monocot, although there are some indications that some dicot promoters can drive gene expression in monocots, such as GUS expression driven by the potato pin2 promoter in rice (Xu *et al.*, 1993). Therefore, MePAL promoter-GUS constructs were also used to assess the promoter activities in rice embryogenic calli and transgenic rice by transient assays and stable transformation, respectively.

5.2 GUS expression driven by MePAL2 promoters in transgenic cassava plants

In most cases of successful cassava transformation, the cultivar TMS 60444 has been used as the target tissue for transgene insertion, due to the ease with which friable embryogenic callus and suspension cultures can be established and plants can be regenerated (Taylor *et al.*, 1996). Therefore, cultivar TMS 60444 was used and Taylor's transformation protocol was followed in the analysis of MePAL2 promoter expression, using stable transformation. In order to study the activity of the PAL promoter during development and the stress response of PPD, and also the effect of deletion of some motifs on the PAL promoter activity, constructs PAL840GUS/MON, PAL400GUS/MON were bombarded into embryogenic callus of TMS 60444 to obtain transgenic cassava plants. Constructs PAL260GUS/pUC and PAL200GUS/pUC were also co-bombarded into the callus respectively, together with pMON977 (see figure 2.1) as selection marker as there was no plant selection marker (*nptII* gene) in these two promoter constructs. Transformation of cassava with CaMV 35 S-GUS as a positive control and promoterless-GUS as a negative

control were also attempted initially but failed and this was not pursued further due to limited resources and the lack of sufficient time for the analysis of transgenic lines.

5.2.1 Identification of positive transgenic lines at the callus stage

The transformed embryogenic calli were selected on paromomycin medium for ten days, then yellowish, large single units of callus were chosen and transferred to freshly prepared selection medium of the same composition. From then, every week (for several weeks), large, yellowish, single colonies were picked from the original plates for the next round of selection. Three weeks after the second round of selection, yellowish single colonies were picked for the third round of selection. Yellowish callus that sustained 3 rounds of selection was considered as an independent paromomycin-resistant transgenic line or paromomycin-resistant callus line.

From 4 experiments, 32 independent resistant callus (or lines) of PAL840GUS/MON were obtained (Table 5.1). Among these lines, 10 lines were contaminated during the move of ILTAB from San Diego to St. Louis (USA) before GUS assays were carried out. The remaining 22 lines were checked for GUS activity and all of them turned blue after staining with GUS staining buffer (X-gluc staining buffer). These lines were named GUS positive lines. In order to confirm that in these lines, GUS was driven by MePAL2 promoters, PCR was carried out to test whether the MePAL2 promoter was located upstream of the GUS gene or not. Before the PCR test, two of the 22 GUS positive lines did not survive paromomycin selection and died. Another two lines were contaminated before the PCR test. From the PAL840GUS transgene, specific primers were designed to cover the promoter region and part of the GUS gene. The forward primer for the PCR test was from 5' end of MePAL2 promoter (PTACF: CACTCGAGTG GAAGCTTATTGTG -925 to –903) (the same primer as used for TATA box mutation construct, section 4.3.3) and the reverse primer was from the 5' end of GUS gene (GUSR: CGTCGAGTTTTTTGATTTCACG). PCR test on the genomic DNA of the rest 18 GUS positive transgenic lines using these primers revealed that all 18 lines showed the expected

PCR products (Table 5.1). The lines showing expected PCR products were considered PCR positive lines.

Similar transformation experiments were carried out with PAL400GUS/MON, PAL260GUS/pUC and PAL200GUS/pUC - transformed lines at the callus stage (Table 5.1). For PAL400GUS/MON, 38 callus lines were resistant to paromomycin over three rounds of selection medium. Among these lines, 5 lines died during the transfer from San Diego to St. Louis before the GUS assay was carried out. The remaining lines, 10 out of 33 lines showed blue after GUS staining, and showed the same GUS expression pattern as PAL840 (Figure 5.1). PCR tests were carried out on 17 of the 33 surviving lines, which including 8 GUS positive lines and 9 paromomycin- but GUS negative lines. The forward primer was PD2 (-247 to –226) TCAACCAACCAAGCTTCTCTCC (the same as used to clone the truncated PAL260GUS promoter construct, section 4.3.2) whose sequence was contained in the PAL400GUS construct, and the reverse primer was GUSR as shown above. The PCR test showed that the eight GUS positive lines contained the PAL400GUS transgene (showed expected PCR products in agarose gel), whereas no PCR products were detected from the other lines (GUS negative).

For construct PAL260GUS/pUC, co-bombarded with pMON977, 8 lines were obtained after three rounds of selection in paromomycin medium (Table 5.1). None of these lines showed blue after GUS staining. PCR tests of these lines using primer GUSR and primer PD3 (-219 to –192, CCTCCTTAAGCTTACCTACCCATTACAC), which was initially designed to clone the truncated PAL200GUS promoter construct (section 4.3.2) showed that 2 of them contained the PAL260GUS transgene and the other lines showed no PCR products.

For construct PAL200GUS/pUC, co-bombarded with pMON977, 45 lines survived after three rounds of selection in paromomycin medium (Table 5.1). Callus of these lines did not show any visible blue spots after GUS staining. However, 5 out of 22 lines tested by PCR (reverse primer, GUSR and forward primer PD4 (-210 to –198)

GCTTACCTACCCATTACAC) did show the expected product from PAL200GUS. Table 5.1 lists putative transgenic lines of different PAL-GUS constructs and PCR test results.

Promoter	Paromomycin resistant callus lines	GUS positive lines (+/tested)*	PCR positive lines (+/tested)*
PAL840	32	22/22	18/18
PAL400	38	10/33	8/17
PAL260	8	0/8	2/8
PAL200	45	0/45	5/22

Table 5.1 Recovery and analysis of putative transgenic callus lines expressing GUS * +/tested indicates the number of lines showing positive among the lines tested.

5.2.2 GUS activities in the early stages of transgenic cassava

The GUS positive lines of PAL840GUS/MON and PAL400GUS/MON were transferred to regeneration medium: Murashige and Skoog (MS) medium supplied with 5 μM α-naphthalene acetic acid (NAA) to induce cotyledons. The PCR positive lines of PAL260GUS/pUC and PAL200GUS/pUC were also transferred to the regeneration medium. After two to three weeks, these lines were transferred to MS medium supplied with 1 μM of NAA. If they still did not reach the cotyledon stage after another two to three weeks then they were transferred again to MS medium supplied with 0.1 μM of NAA (section 2.7). The rate of regeneration for individual lines varied. Line 2 of PAL840GUS/MON was the first line that regenerated cotyledon, followed by line 7 and line 13 of PAL840GUS/MON. But some lines, such as line 3 and line 5 of PAL840GUS/MON, did not respond to the regeneration medium and no cotyledons were induced even after three months in MS medium.

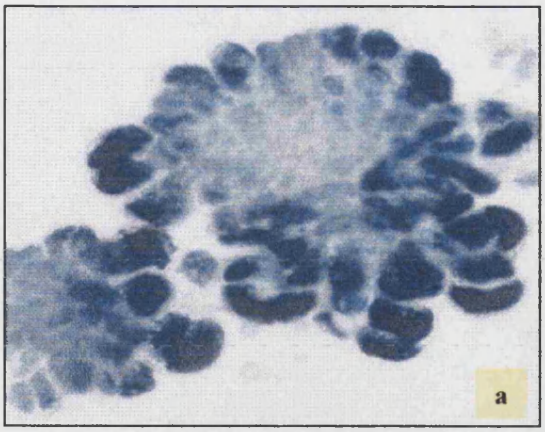
Once cotyledons were formed, they were transferred to MS2 2BAP (section 2.7) to induce shoots. Shoots were induced in most lines with cotyledons (Table 5.2). The shoots were then transferred to MS medium to induce rooting. After a month on MS medium, roots were induced and small *in vitro* plantlets were obtained, which were then transferred to compost and grown in the greenhouse (section 2.7)

Constructs	Callus lines	Cotyledon stage	Shoot induction	Root induction
PAL840GUS/MON	22	11	11	11
PAL400GUS/MON	10	4	4	4
PAL260GUS/pUC	2	1	0	0
PAL200GUS/pUC	5	0	0	0

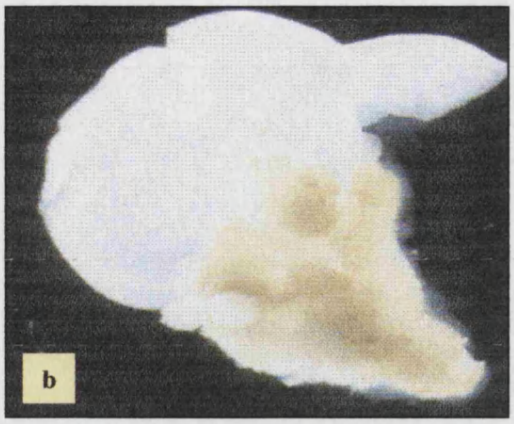
Table 5.2 Transgenic lines obtained at different stages.

GUS activity was investigated in this early stage (cotyledon) of these cassava transgenic lines. During the callus stage on selection medium, GUS was strongly expressed in PAL840GUS/MON lines, but then in most lines, such as line 2, GUS expression decreased at torpedo and cotyledon stages on regeneration medium (Figure 5.1). Some lines, such as line 7, decreased such that there was no visible GUS blue stain at the cotyledon stage. The only exception was line 10, in which GUS expression stayed very high until the cotyledon stage (Figure 5.1). All four lines of PAL400GUS/MON showed the same pattern of GUS expression as did most lines of PAL840GUS/MON: strong expression in callus and decreased expression in torpedo and cotyledons.

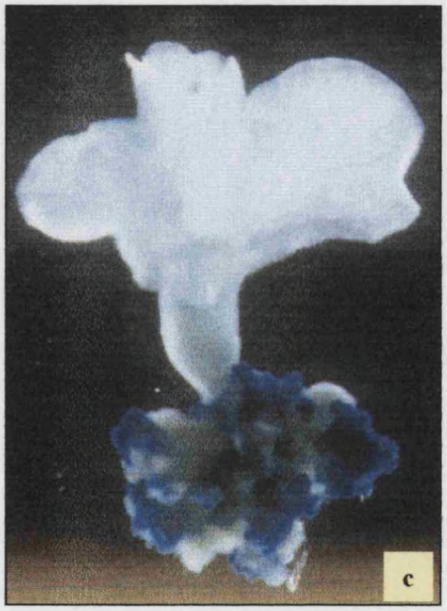
It was unfortunate that after shoot induction there were not any transgenic cassava plants from constructs PAL260 and PAL200. This may partly be due to the limited number of transgenic callus lines obtained to start with for shoot induction. Four lines of PAL400 were regenerated, but no GUS activity test was done as the plantlets were too young to be



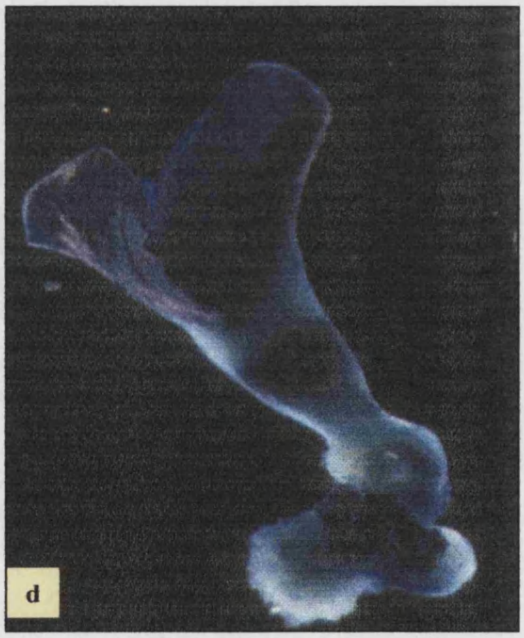
a: GUS expression in the transgenic cassava callus (line 2). The pattern is the same in all the lines of PAL840GUS/MON.



b: Negative control (cotyledon of wild type TM 60444) for GUS staining.



c: GUS expression in the transgenic cassava cotyledon (line 2). All the PAL840GUS/MON lines except line10, showed this GUS expression pattern. GUS expression is very weak or not visible at the cotyledon stage.



d: Strong GUS expression in cotyledon of transgenic cassava PAL840GUS/MON line 10.

Figure 5.1 GUS expression of PAL840GUS/MON from callus to cotyledon stage in transgenic cassava.

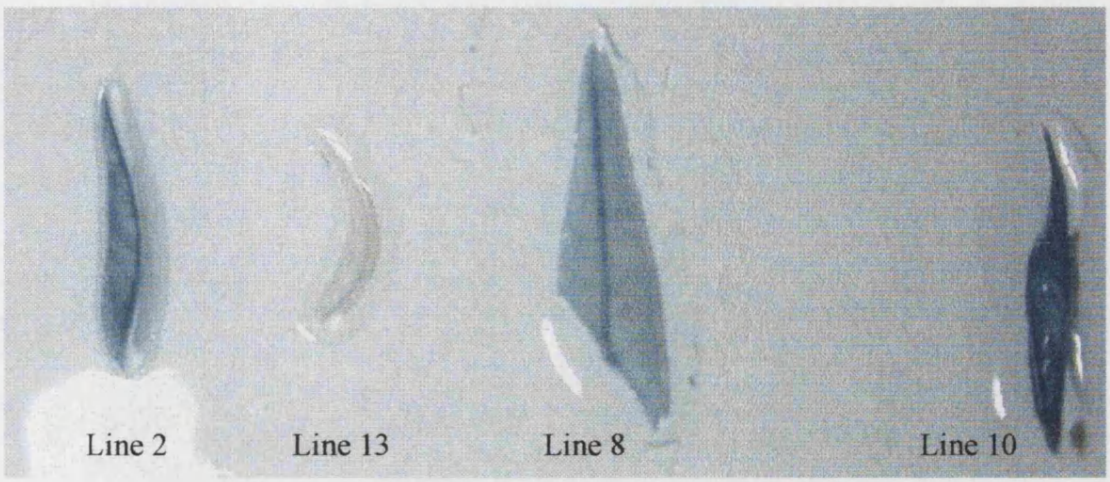
GUS staining condition: at 37°C in the dark for 2 hrs for callus (a) and 16 hrs for cotyledons (b, c and d).

excised for the test before experimental work ceased. Therefore, in the following sections, the GUS analyses focus on PAL840GUS/MON transgenic cassava lines.

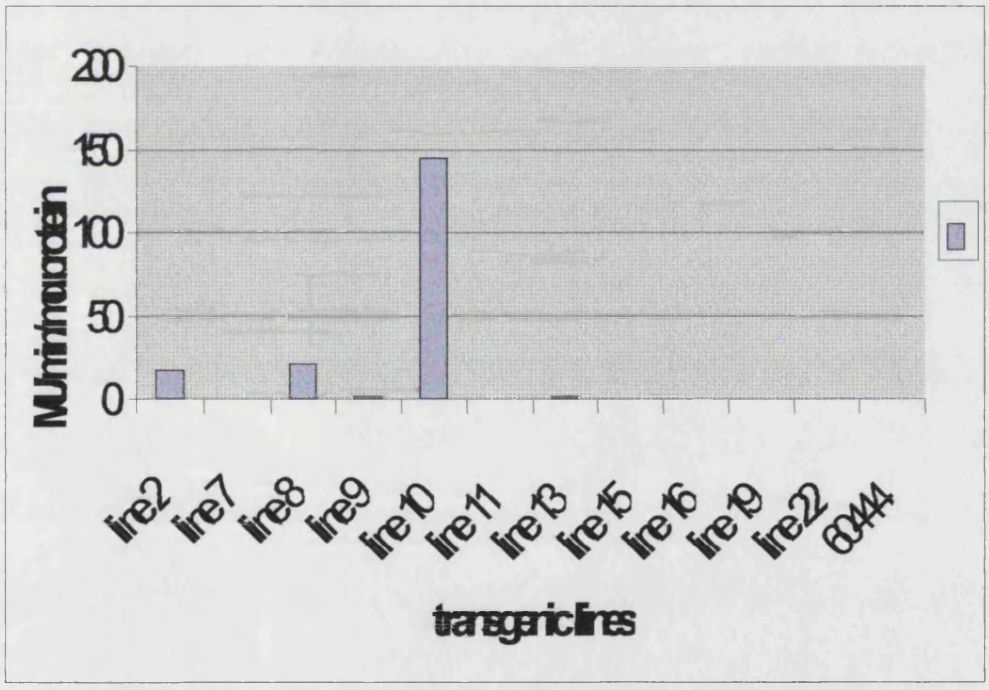
5.2.3 GUS activities directed by PAL840 in transgenic cassava during development

In other plant systems, PAL genes have been shown to be developmentally regulated (Gray-Mitsumune *et al.*, 1999). In order to test whether this was the case for MePAL2, GUS activity was investigated in every stage during the development of PAL840GUS/MON cassava transgenic lines. At the plantlet stage, lines 7, 11, 15, 16, 19 and 22 of PAL840GUS/MON, which were GUS positive at the callus stage, did not show GUS blue when stained (Figure 5.2b). GUS activity was extremely low (detected by MUG assay) (Figure 5.2b) in line 9 and it did not show visible GUS blue staining. Lines 2, 8, 10 and 13 showed GUS blue in the vascular tissue of leaves, stem and roots, and root tips (Figure 5.3). However, the GUS activities of these lines varied as shown by histochemical GUS staining in leaves (Figure 5.2a) and MUG assays on the leaves (Figure 5.2b). Line 10, showing strong GUS expression at the cotyledon stage, also showed strong GUS expression at the plantlet stage (Figure 5.2). In line 13, GUS activity was just enough (very low) for visible GUS blue in the vascular tissue of a leaf (Figure 5.2). The variation of GUS expression among transformed plant lines has been observed in a number of other plant systems (Barfield and Pua, 1991; Ishige *at al.*, 1991; Kemo *et al.*, 2000; Gittins *et al.*, 2001).

In order to test whether MePAL2 was developmentally regulated, GUS activities were quantified in different stages of different organs from six-month old transgenic cassava plants grown in the greenhouse. Line 2 was the first regenerated GUS-positive line with a number of propagated plants while other lines were still at the regeneration stage or the early stage of plantlets. Therefore, a plant of Line 2 grown in the greenhouse, 1.7 meter high and with 44 expanded leaves (see Figure 5.7), was used for the GUS activity assays. Stem sections, petiole and leaves were harvested every five nodes down from the apex to the base of the stem, which represent the different stages of the development of these



a: Histological analysis of GUS expression in the *in vitro* leaves of different lines.



b: MUG analysis of GUS activities in the *in vitro* leaves of different lines

Figure 5.2 GUS expression level in the *in vitro* leaves of different transgenic cassava lines containing PAL840GUS/MON.

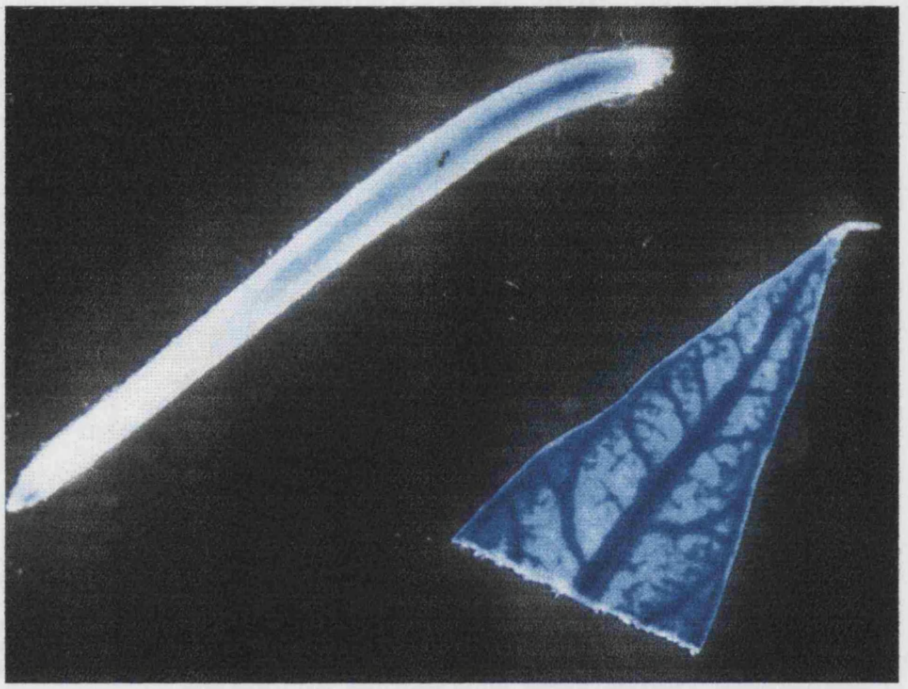


Figure 5.3 GUS expression in root and leaf of *in vitro* transgenic cassava plant (line 2).

GUS was expressed in root tip and vascular vessels of root. In leaf, GUS was expressed in vascular tissue.

organs from growth to senescence. GUS activities in newly opened leaves and their petioles and attached young stems were at similar levels, about 5 nmol MU/min/mg protein (Figure 5.4). From the 5th node downwards, GUS activities in leaves were 3 fold to more than 50 fold less than in their petioles and stems (Figure 5.4). GUS expression in young leaves was slightly higher than in old leaves. GUS activities increased dramatically from petioles of young leaves to mature petioles (around 20th node) and then tended to decrease with senescence (Figure 5.4). GUS activities in stems of different positions, from top to base, seemed to follow a pattern similar to petioles but the peak activity appeared in the stem section 5 nodes further down (Figure 5.4).

5.2.4 GUS expression patterns in different tissues of transgenic cassava

To determine the cell- and tissue-specific patterns of GUS expression driven by PAL840, leaves, petioles and stems harvested from the top, middle and lower part of a six-month old cassava plant (Line 2) grown in the green-house were free-hand sectioned and then the GUS activity was histochemically assayed. Roots including fibrous roots, thickening roots and tuberous roots were also sectioned and analysed.

In the leaf, GUS was strongly expressed in the midrib or vascular tissue (Figure 5.5a). Leaf sections showed that GUS expression was localised in xylem parenchyma (Figure 5.5b).

GUS expression patterns were also tested in different parts of the petiole including near-leaf, middle of the petiole and near-stem part. GUS staining showed that expression was localized to the xylem parenchyma in the vascular system (Figure 5.6).

Stem sections were taken from the top (young stem near to the apex), middle and lower part (highly-wooded stem) of the plant. GUS activity was localized to the xylem parenchyma of the stems from different parts of the plant (Figure 5.7). It was also noted that in the woody stem, GUS was also expressed in the epidermis and bark cambium

(Figure 5.7). Interestingly, tyloses were found in the stem and GUS was strongly expressed in the tyloses of transgenic cassava stems (Figure 5.8).

In cassava fibrous roots, roots that were undergoing thickening, and tubers, the GUS was expressed in the xylem parenchyma (Figure 5.9). GUS was also expressed in the cambium of tubers. GUS was strongly expressed in tyloses of thickening roots.

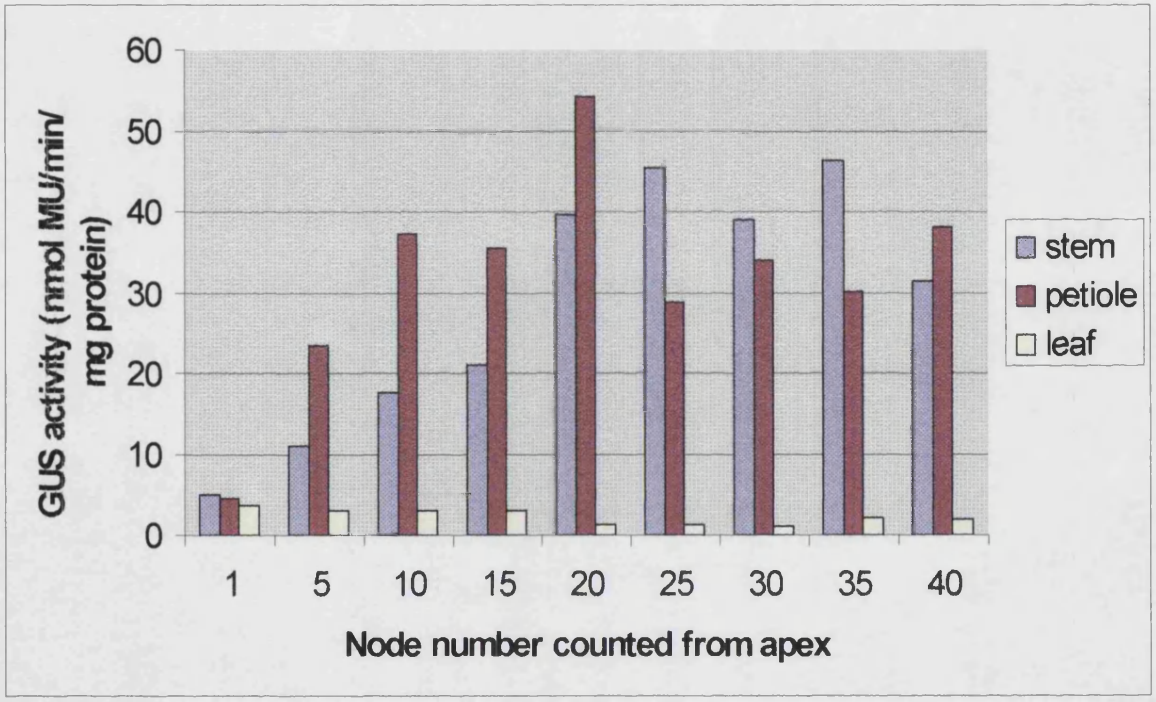


Figure 5.4 GUS activity in leaves, stems and petioles at different stages of development in a transgenic cassava plant Line 2 containing PAL840GUS/MON.
Stem, petiole and leaf were harvested every five nodes down from the apex to the base of the stem, which represents the different stages of the development of these organs from growth to senescence.

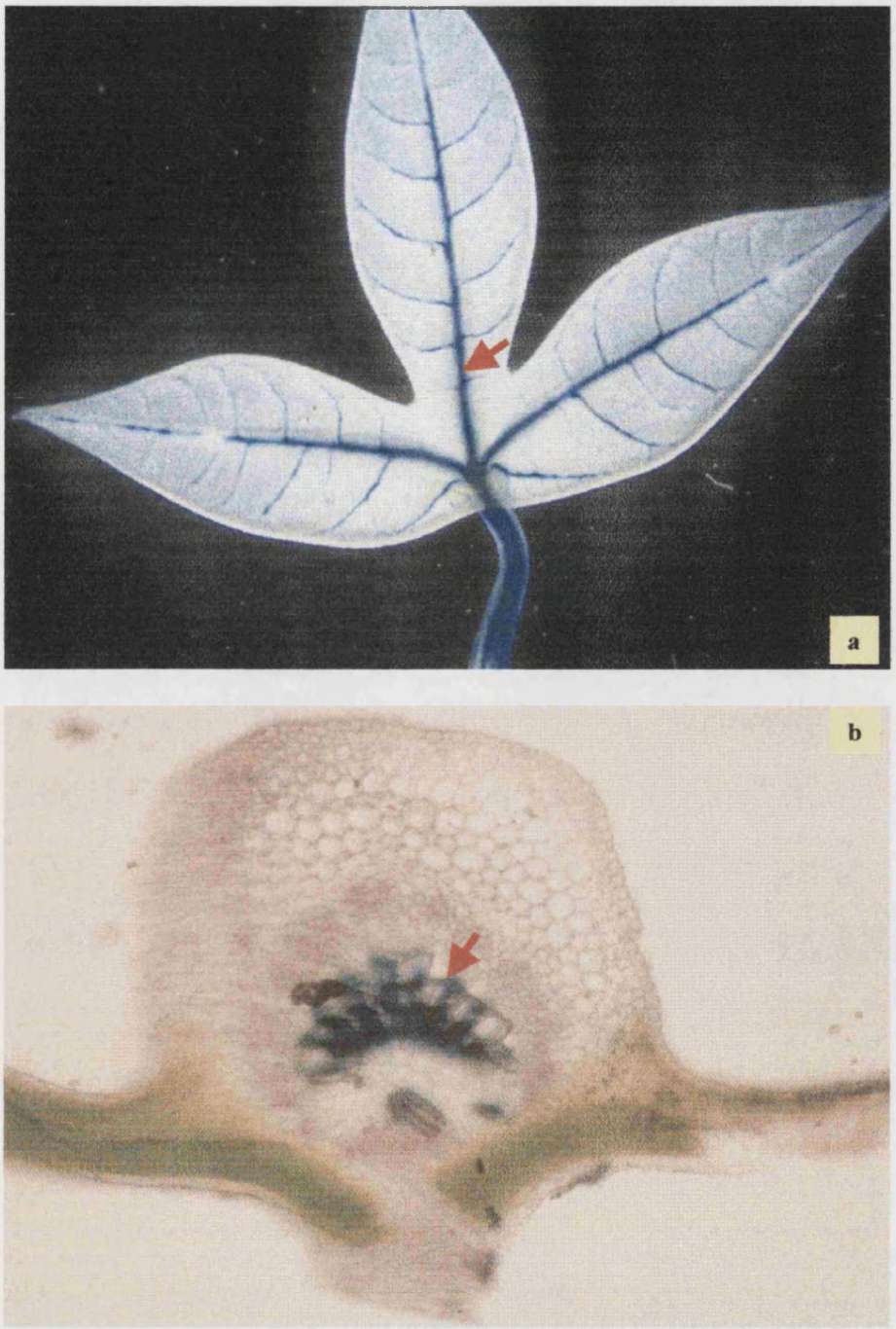
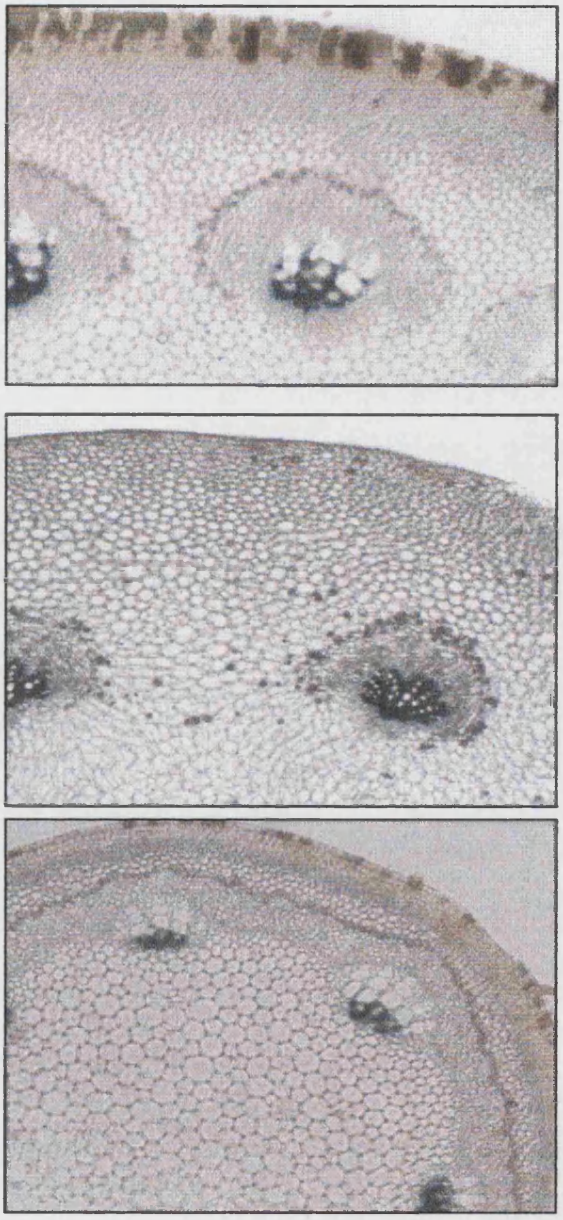


Figure 5.5 GUS expression in a cassava leaf from transgenic line 2 containing PAL840GUS/MON.
a: An young transgenic cassava leaf was stained with GUS staining buffer overnight, at 37°C in the dark.
b: Free-hand cross-section of transgenic cassava leaf and its GUS expression pattern.
Red arrows indicate the location of GUS expression.



The position of the section is the end of petiole that near a leaf

The position of the section is at a middle of the petiole

The position of the section is the end of petiole that near a stem

Figure 5.6 GUS expression in a cassava petiole from transgenic Line 2 containing PAL840GUS/MON

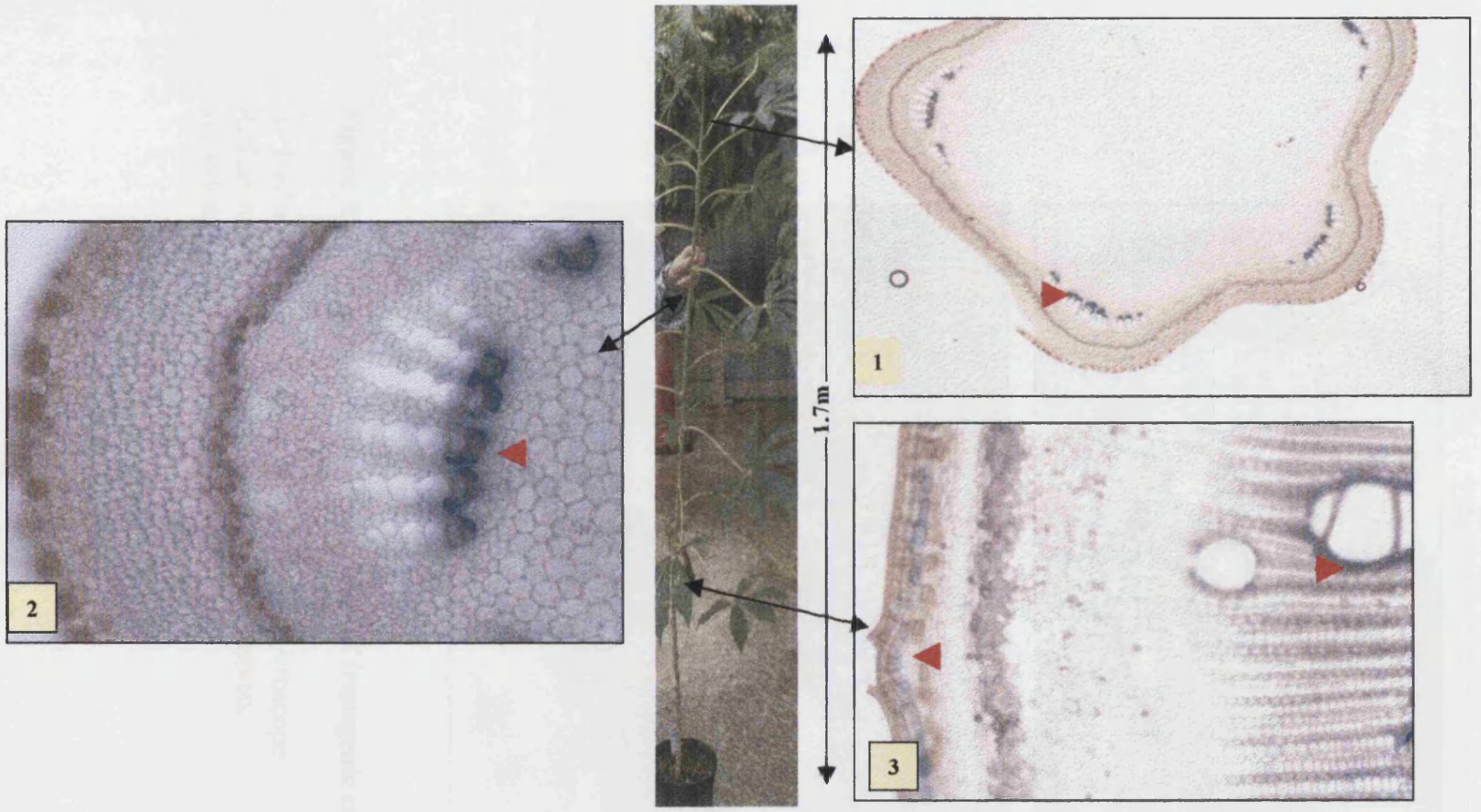


Figure 5.7 GUS expression in the stem of transgenic cassava (line 2 of PAL840GUS/MON).

The plant line 2 was analysed when it was 1.7meter high, 6 month old (as shown in the middle). Black lines show the positions from where the sections were taken. Red arrows indicate the location of GUS expression.

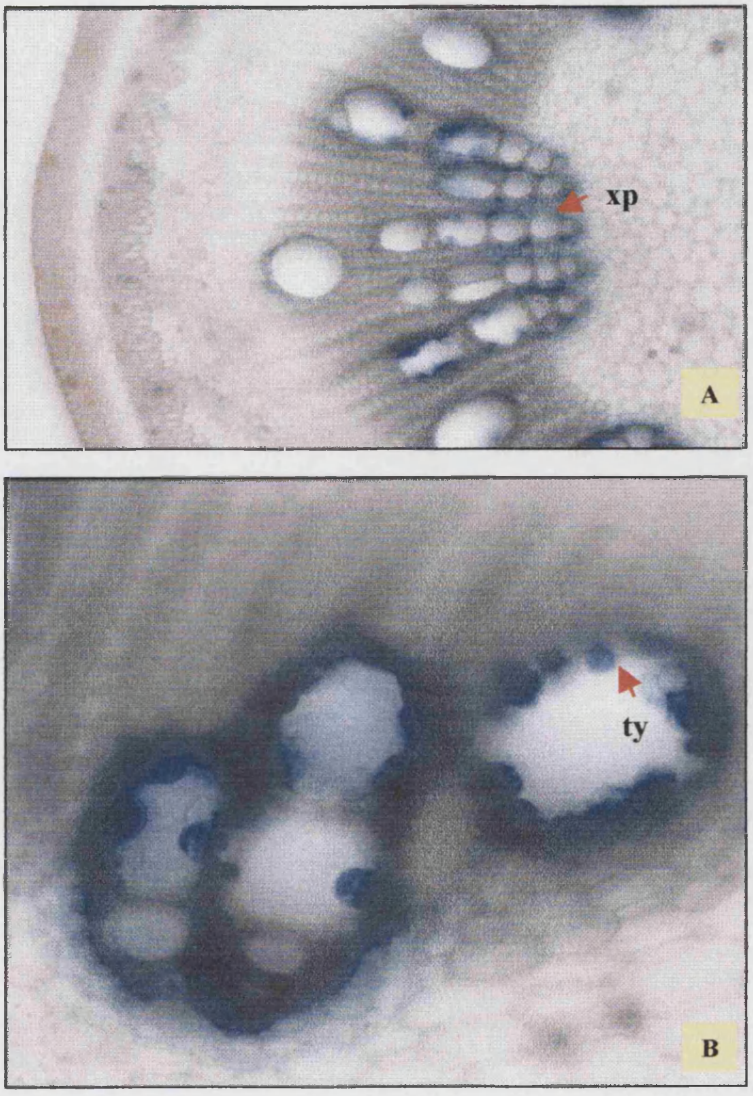


Figure 5.8 GUS expression in stem tyloses of transgenic cassava.

A: Under 20x of microscope B: Under 40x microscope
Red arrows indicate the location of GUS expression.
xp, xylem parenchyma cells, **ty,** tyloses

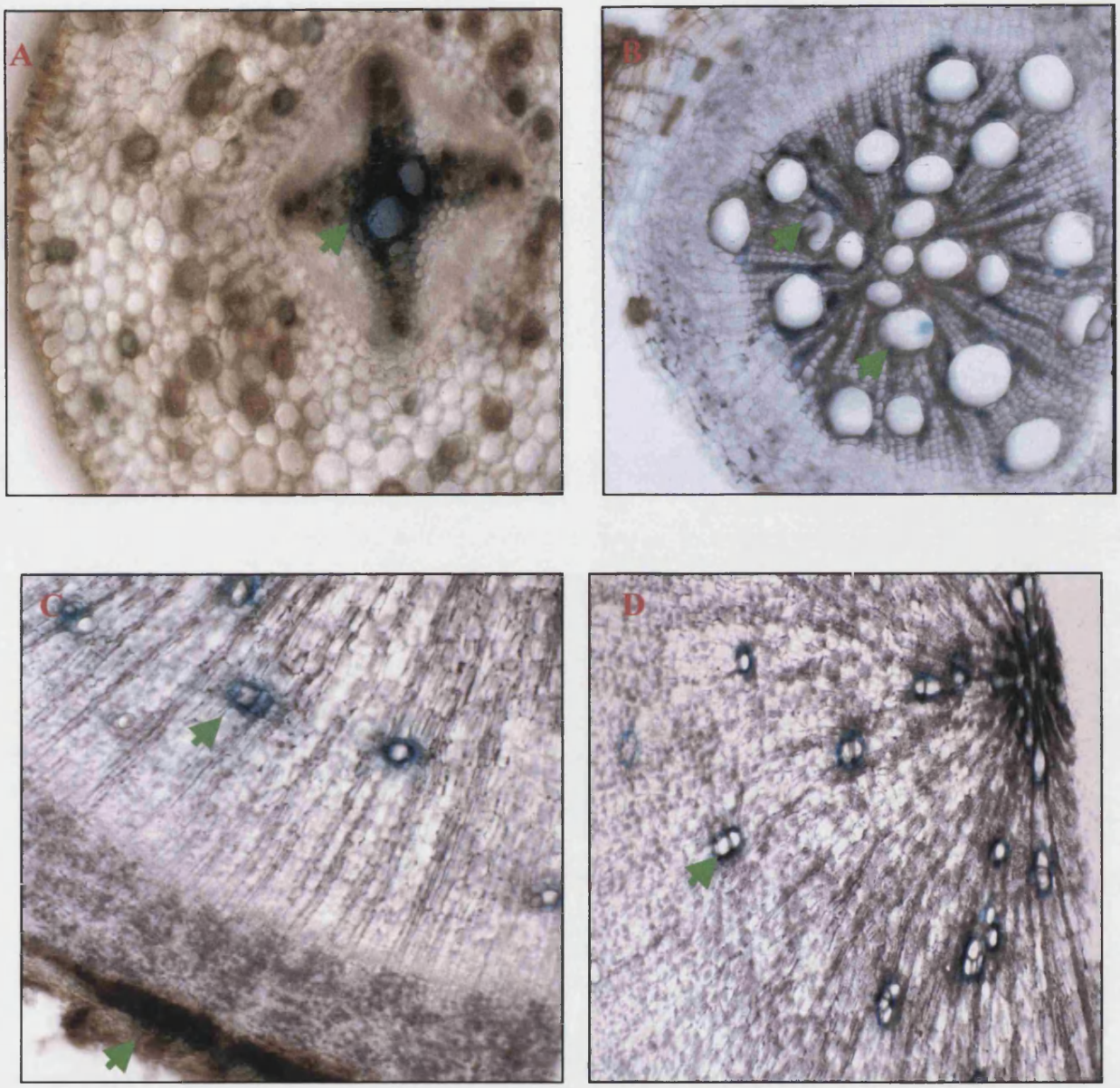


Figure 5.9 GUS expression in cassava vegetative roots and storage roots.

A: Vegetative root or fibrous root **B:** Root undergoing thickening. **C & D:** Cassava storage roots. The location of GUS activity is indicated by green arrow.

5.2.5 GUS expression in response to PPD

Post-harvest physiological deterioration (PPD) is a main constraint in cassava storage. It seriously affects the marketability and palatability of cassava. In most cassava cultivars, PPD can occur within 24 hours after harvest. The initial symptoms of PPD are characterized by vascular or xylem vessel discolouration (streaking), and then followed by general discolouration of the storage parenchyma. It is believed that PPD could be a sustained wound reaction spreading systemically from the wound site into the whole root (Beeching *et al.*, 1994). During PPD, increased phenolic compounds via the phenylpropanoid pathway are produced. PAL is the first and key entry enzyme to the core reaction of pheylpropanoid metabolism. Therefore, the properties of MePAL2 promoter during cassava PPD were investigated.

As shown in section 5.2.3, GUS was expressed in the xylem parenchyma of tuberous root, where the visible symptoms or discoloration occur during PPD. A transgenic tuber was harvested from a plant of Line 2 and incubated at room temperature in the dark for two days. Part of the tuber was transversely sectioned and the other part was cut along the tuber. The cross-section and longitudinally cut tuber sections showed typical PPD symptoms (Figure 5.10a and b). These sections were then stained for GUS expression patterns, which showed that GUS activity was localised to vascular bundles, exactly the same location as the blue-brown discoloration (Figure 5.10c and d). Another question to ask is how the activity of MePAL2 promoter behaves during PPD. Firstly, the GUS expression patterns of tubers immediately after harvest, and tubers incubated for 1 day, 2 days to 4 days after wounding, were histochemically analysed. In detail, fresh-harvested tuber was cut (artificial mechanical wounding) transversely into 4 cm thick slices, and kept in the dark at room temperature and harvested every 24 hours. The samples were then hand-free sectioned and stained with GUS staining buffer. GUS was expressed in xylem parenchyma in all the sections taken from different stages of PPD and there was no visible difference in the expression pattern between these sections.

Secondly, GUS activity of tubers undergoing PPD was quantified by MUG assays. Cassava tuber samples were collected three months and six months after propagation from cuttings. The tubers harvested after three months were about 2 cm in diameter and 4 cm in diameter by six months. The harvested tubers were cut transversely into 2 cm thick slices and kept in the dark at room temperature for 24 hours, 48 hours, 72 hours and 96 hours. GUS activities of the protein extracts from these samples were measured by MU (Perkin-Elmer LS). Tubers from cultivar TM60444 harvested and treated in the same way as the transgenic Line 2 were used as controls. In the tuber harvested 3 months after propagation, GUS activity increased significantly 24 hours after wounding, declined at 48 hours (still one fold higher than 0 hour) and then increased to a level similar to 24 hours. In the tuber harvested 6 months after propagation, GUS activity increased steadily to a peak 72 hours after wounding and then declined (Figure 5.11).

Pruning, or cutting off the top of cassava 20-30 cm above the base of the stem 2 to 3 weeks before harvest, can reduce the susceptibility to PPD (Tanaka *et al.,* 1984). A plant of Line 2 (PAL840GUS/MON) was pruned two weeks before harvest and then GUS expression pattern and activities were analysed. The GUS expression pattern during PPD in the tuber from the pruned plant was the same as that of the tuber from the non-pruned plant (Figure 5.12a and b). However, it is interesting to note the difference between the tuber sections from un-pruned plants and pruned plants, with the latter being nearly transparent. Based on the increased soluble sugar contents in the tubers of pruned plants (Data *et al.,* 1984; Hirose, 1986), Han (2000) suggested that pruning leads to the switch of tubers from being as a sink to a source, which can increase the water retaining ability of the tuber and delay PPD. Therefore, the transparent background of the tuber section from pruned plants is probably due to the decomposition of starch to soluble sugars. Analysis of GUS activities of the tubers from the pruned plant during PPD by MUG assays revealed that GUS activity increased to a peak 48 hours after harvest, which was later than the timing (24 hours) of the appearance of the peak GUS activity in tuber from non-pruned plants (Figure 5.12c). However, the activities in tubers from pruned and non-pruned plant were similar 48 hours after harvest (Figure 5.12c).

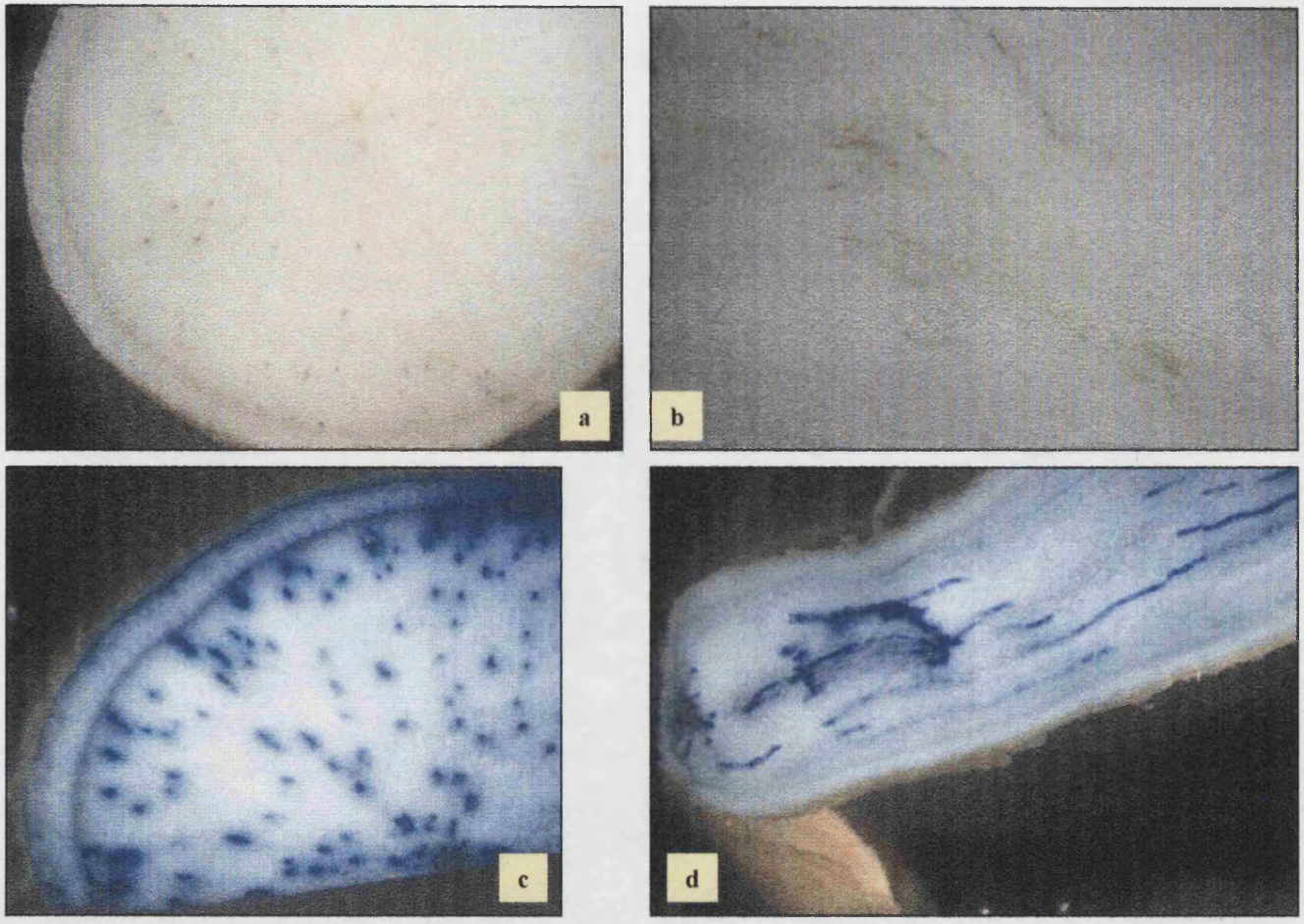


Figure 5.10 GUS expression in cassava tuber in response to PPD.

a and b: PPD symptoms-cassava tubers were incubated at room temperature in dark for 2 days after harvest and sectioned (a) transverse section (b) longitudinal section. **c and d**: GUS expression pattern of cassava tubers during PPD. The sections in a and b were stained with GUS staining buffer.

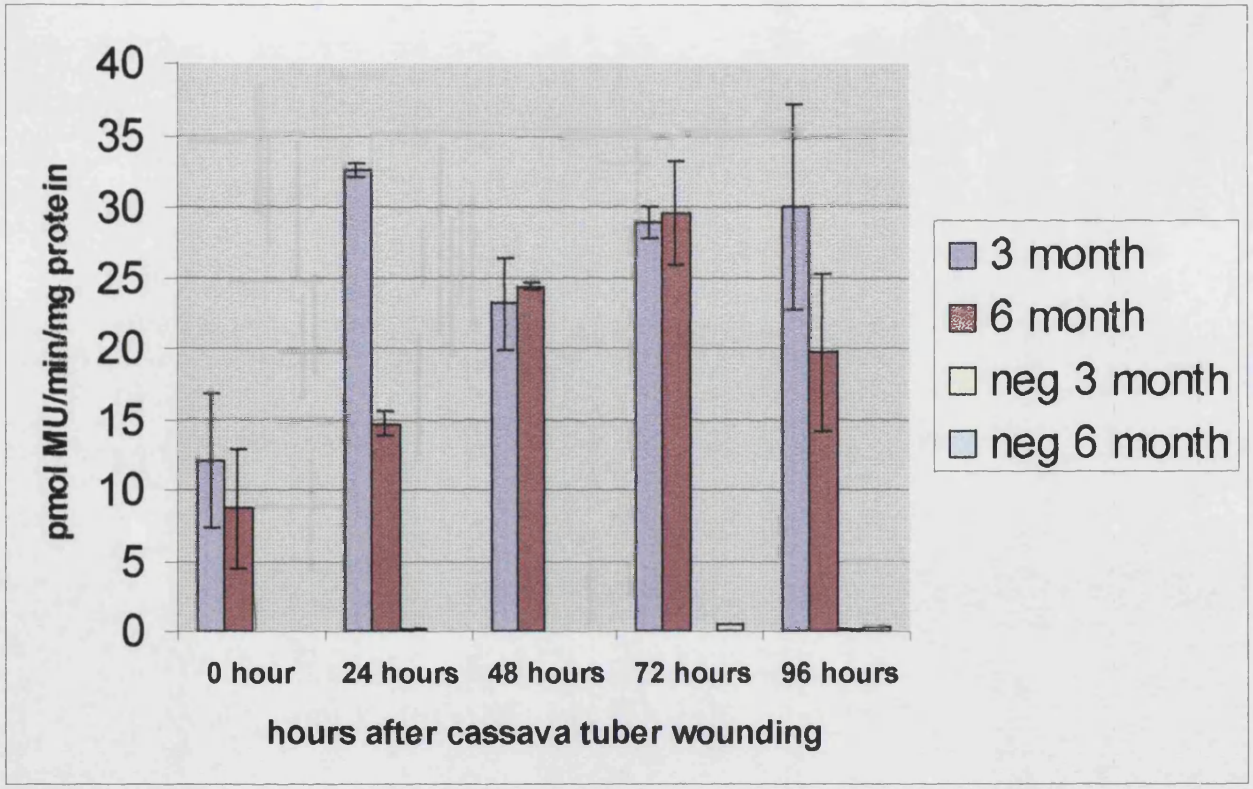
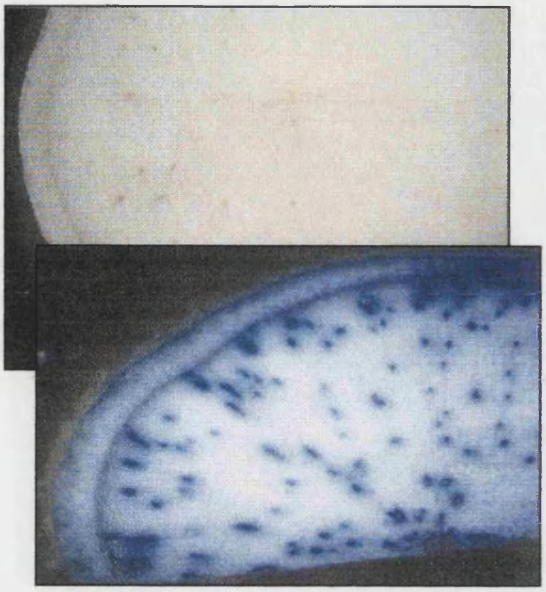
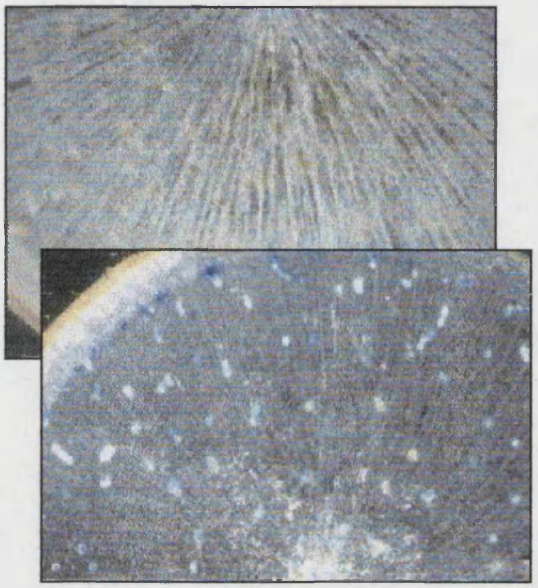


Figure 5.11 GUS activities of MePAL840GUS in response to PPD.

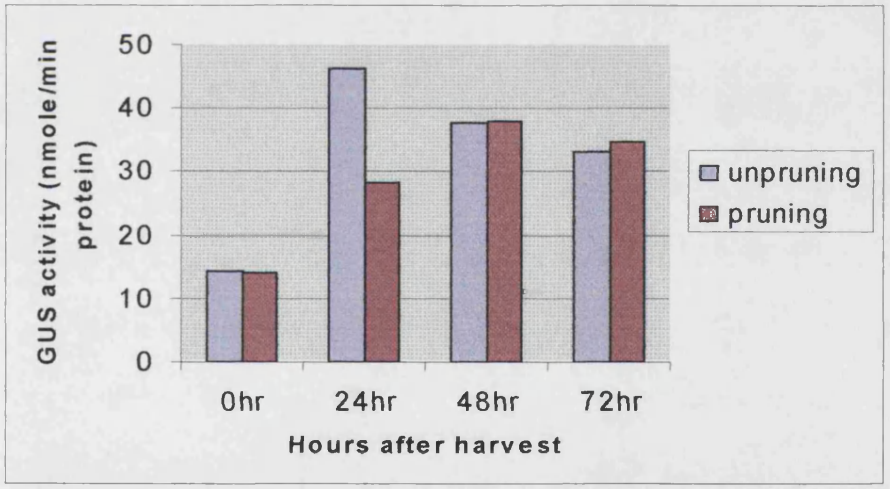
MePAL840GUS/MON line 2 tubers harvested 3 and 6 months after propagation were cut into 2cm slices, incubated at room temperature in the dark, harvested every 24 hrs and extracted for proteins for MUG assay. Tubers from TM 60444 were used as negative control.



a: Vascular discoloration (upper) and GUS staining (lower) of deteriorating tuber (2 days post-harvest) from un-pruned plant.



b: Vascular discoloration (upper) and GUS staining (lower) of deteriorating tuber (2 days post-harvest) from pruned plant.



c: MUG analysis of tubers from pruned and un-pruned plants during PPD.

Figure 5.12 The effects of pruning two weeks before harvest on MePAL840 promoter activity during physiological deterioration (measured by GUS activity).

5.3 Expression patterns of MePAL2 promoter in transgenic rice plants

All studies on PAL promoters using GUS fusions to date were carried out in either homologous hosts (dicot or monocot or gymnosperm) or the model plant, tobacco. The expression patterns of the PAL-promoter-GUS construct were usually the same in the heterologous host as in the homologous host. One exception is poplar PAL1 and PAL2, GUS expression driven by poplar PAL1 or PAL2 in transgenic tobacco did not accurately reflect PAL2 expression in poplar (Gray-Mitsumume, 1999). Contrary to PAL1/2 expression patterns in poplar, no expression of either fusion (PAL1-GUS and PAL2-GUS) was detected in epidermal or subepidermal cell layers of young tobacco leaves or stems. However, until now, no study has been done on the expression of a dicot PAL promoter GUS fusion in a monocot. In order to investigate the expression pattern of the MePAL2 promoter in a monocot, truncated MePAL2 promoter constructs, PAL840GUS/pUC, PAL400GUS/MON, PAL260GUS/pUC and PAL200GUS/pUC were separately co-transferred with a hygromycin resistant gene plasmid (p35H) (see Figure 2.1) into rice callus to test the stable activity of the PAL promoter.

Before stable transformation was carried out, PAL840GUS/pUC was bombarded into rice callus to test whether the cassava MePAL2 promoter was active or not. The transient assay of the PAL promoter showed that it could drive strong GUS expression in rice callus (Figure 5.13a).

The protocol used for rice callus induction and gene transformation was adopted from Chen *et al.* (1998). Embryogenic calli were induced from seeds of cultivar TP309, which has been successfully used for rice transformation. The calli were then amplified 3 to 4 times by transferring them to fresh-made callus induction medium until there were enough calli for rice transformation. Different constructs of PAL promoters were then transferred together with p35H into embryogenic calli by particle bombardment. These calli were transferred onto hygromycin selection medium. For PAL840GUS/pUC, after 5 weeks in the hygromycin selection medium, 62 relatively compact and opaque calli from

PAL840GUS/pUC were visually identified and those calli were considered individual hygromycin resistant lines. PAL400GUS/MON, PAL260GUS/pUC and PAL200GUS/pUC, 28, 12 and 10 hygromycin resistant lines were obtained respectively. Histochemical analysis of these resistant lines showed that 32 lines of PAL840GUS/pUC were GUS positive, and Lines 6, 4 and 2 were PCR-positive for PAL400GUS/MON, PAL260GUS/pUC and PAL200GUS/pUC respectively. The GUS expression pattern at this stage (Figure 5.13b) was the same among the different constructs (data not shown).

At the plantlet stage, GUS expression patterns were examined in young leaves. Only 8 lines of PAL840GUS/pUC were GUS positive in the leaf and the rest did not show any visible GUS blue after overnight staining. Three lines of PAL400GUS/MON showed GUS positive but there was no GUS positive line for constructs PAL260GUS/pUC and PAL200GUS/pUC. The GUS expression patterns in partial leaves of PAL840GUS/pUC and PAL400GUS/MON lines were the same, i.e, GUS blue was visible at the site of the cut rather than within the leaf (Figure 5.13c). One reason for the expression of GUS at the cut site and not in the whole leaf, could be that it is difficult for staining solutions to penetrate the rice leaves. It was shown that GUS expression driven by a ubiquitin promoter, a constitutive promoter extensively used in plant transformation, showed the same pattern in rice leaves (Chen *et al.*, 1998). In order to gain detailed GUS expression pattern in leaves, transgenic rice leaves, transformed with PAL840GUS/pUC construct(s) were cross-sectioned and then stained for GUS activity. In sectioned leaves of most lines, GUS was abundantly expressed in vascular bundles, bulliform cells, stomata and trichomes (Figure 5.14a). In some lines, GUS was expressed just in the vascular bundle and trichome. In line 9, GUS was only expressed in tissues linking vascular bundles, similar to the GUS expression pattern in leaf sheath (Figure 5.14c).

Four weeks after the *in vitro* rice plants were moved to a greenhouse, roots, stems, leaves and leaf sheath of these plants were analysed for their GUS expression patterns. Generally, GUS expression patterns in young leaves of most lines were the same as observed in leaves at the *in vitro* stage, i.e., GUS expression in vascular bundles, bulliform cells, stomata and

trichomes. The lines showing strong GUS expression at *in vitro* stage maintained their high level of GUS expression in young leaves and the lines showing weak GUS expression were still so in the young leaves. In the leaf sheath, GUS was expressed in the vascular bundles and vascular linkers (Figure 5.14c). In the stem and node, GUS was also expressed in the vascular bundles (Figure 5.13e, Figure 5.14d and 4.24e). In the root, generally, GUS was expressed in root tips, root hairs and vascular cylinder. In root sections, GUS was expressed in the pericycle, exodermis and lateral root initiation sites (Figure 5.14b). Transgenic rice flowers also showed strong GUS expression in anthers (Figure 5.13d).

Among 32 lines transformed with PAL840GUS showing GUS expression at the callus stage, there were only 8 lines that showed GUS expression at the *in vitro* stage, whereas the other lines showed no visible GUS activity. In order to clarify this point, PCR was used to check if the promoter and GUS gene were present in the transgenic rice plants. A forward primer PD1 was designed from 5' of PAL promoter region and a reverse primer GUSR was designed from *uidA* (GUS) gene. Ten lines showed GUS positive at the callus stage, of which 5 lines (Lines 1, 2, 4, 8 and 9) were GUS positive and the others (Lines 3, 5, 6, 7 and 10) were negative in young leaves, were tested, with wild type rice (cultivar TP309) as a negative control. PCR products with expected size based on the positions of the primers were amplified from 9 lines excluding line 10 (Figure 5.15). This suggests that the Lines 3, 5, 6 and 7 contained the PAL promoter-GUS transgene but the transgene was not expressed.

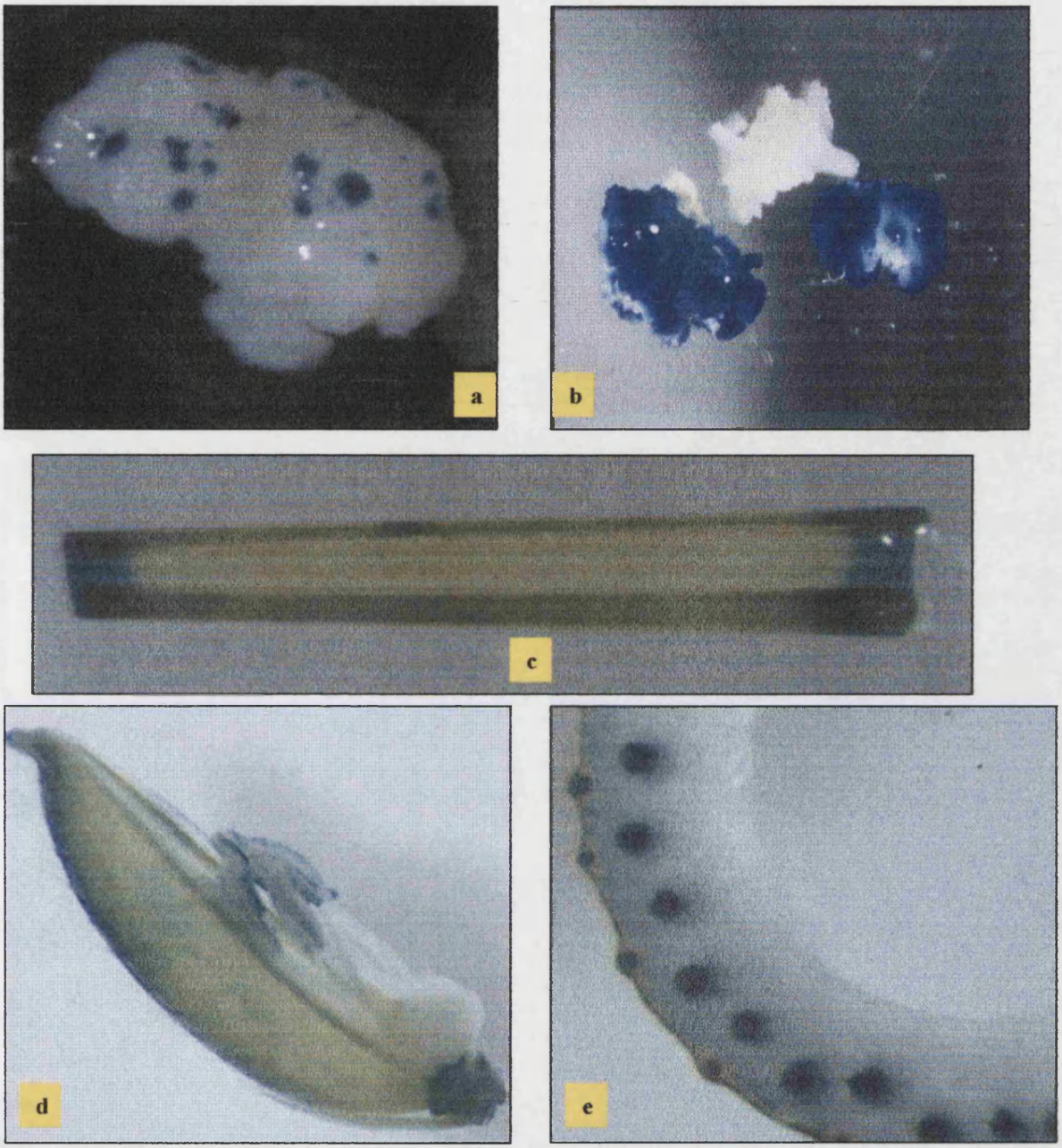


Figure 5.13 Transient and stable GUS expression in rice plant.

(a) Transient GUS activity in rice callus. **(b)** Stable GUS activity in rice callus, callus on the top was negative control and the two below were bombarded with PAL840GUS/pUC. **(c)** Stable GUS expression in rice leaf. **(d)** GUS expression in rice flower. **(e)** GUS expression in rice stem.

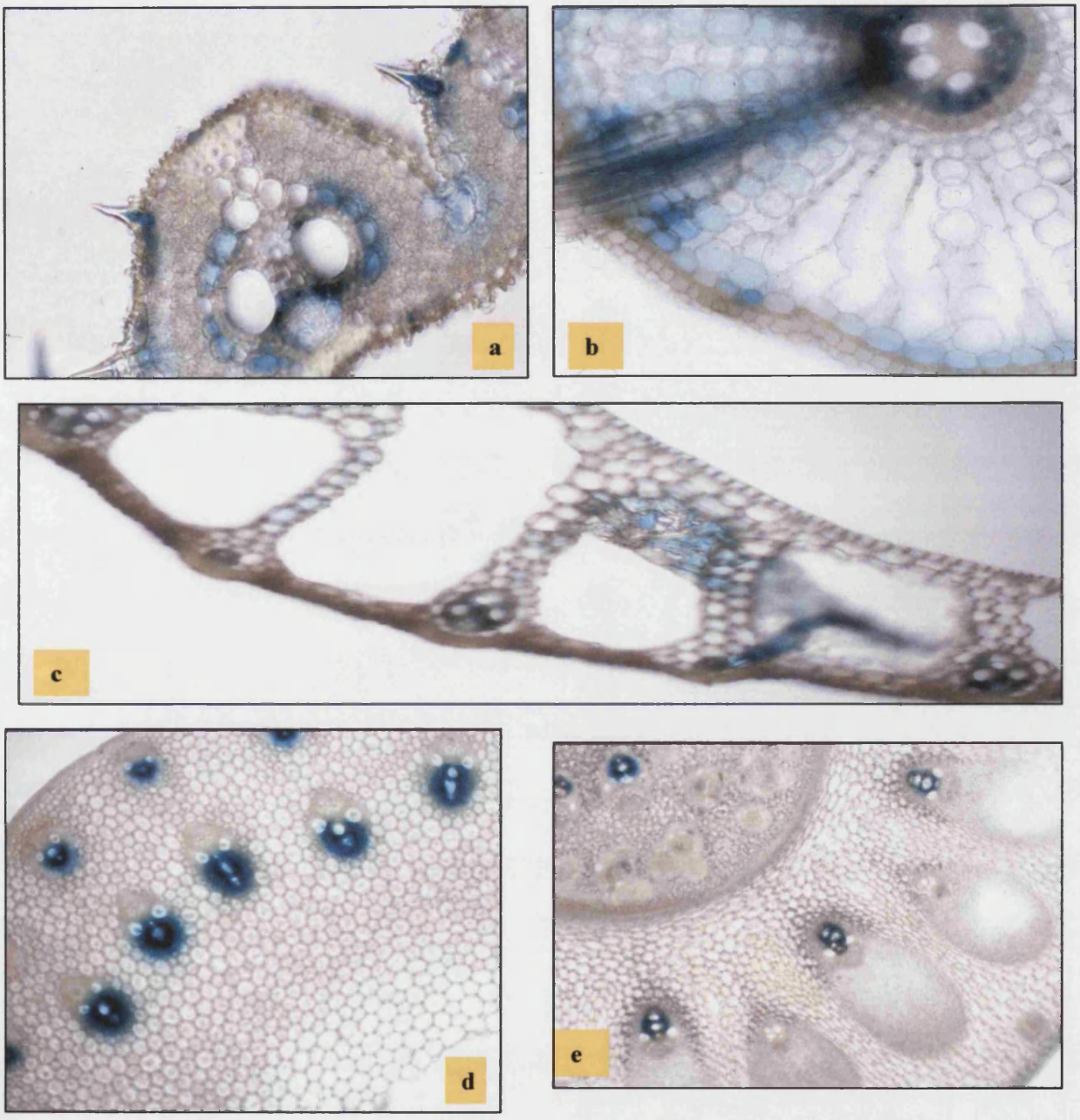


Figure 5.14 Histochemical localization of GUS activity in transgenic rice transformed with PAL840GUS/pUC.

a: young leaf section (20x), **b:** root section (40x), **c:** leaf sheath section (20x), **d:** young stem section (20x), **e:** node section (20x).This data was from line 1.

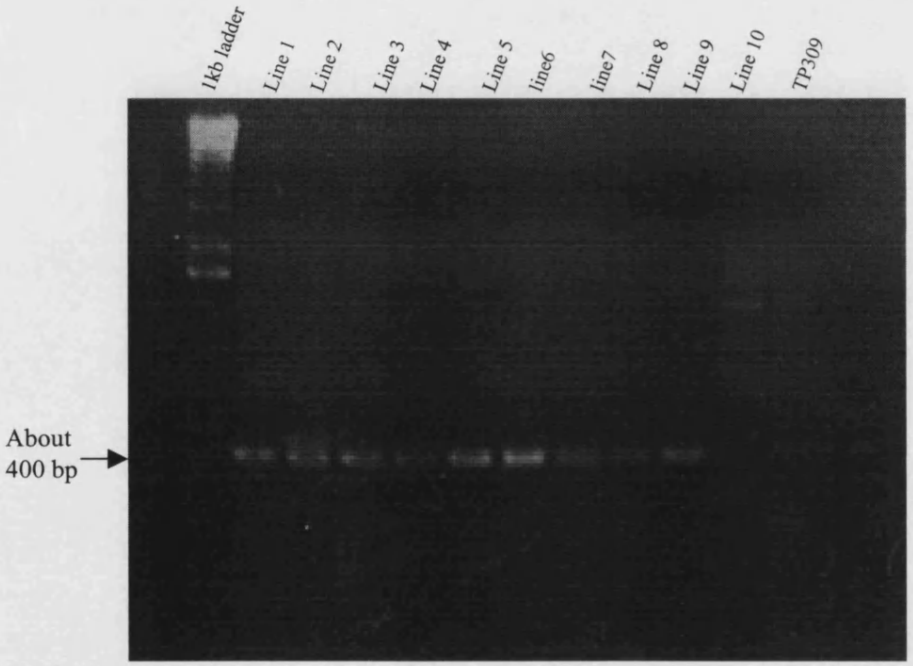


Figure 5.15 Confirmation of transgenic rice lines by PCR.

Genomic DNA extracted from putative transgenic lines were tested by PCR using primers PD1 and GUSR spanning part of MePAL2 promoter and part of GUS gene. PCR conditions were as described in chapter 2.

5.4 Discussion

This is the first report that a cassava promoter successfully drove the expression of a reporter gene in cassava, though there are several cases of reporter gene expression driven by promoters isolated from other organisms such as 35S from cauliflower virus and CsMV from cassava mosic virus (Schopke *et al.,* 1996; Li *et al.,* 1996; Taylor *et al.,* 1998). This research is also the first example of application of cassava transformation technique to cassava gene analysis. In this chapter, the functional properties of the cassava MePAL2 promoter in transgenic cassava and rice plants are described.

5.4.1 Deletion analysis of MePAL2 promoter

MePAL2 promoter segments including PAL840, PAL400, PAL260 and PAL200 (Section 4.3 for construct details) were fused with the GUS gene and transferred into embryogenic cassava calli. Transgenic cassava plants were generated to test the activities of these truncated promoters.

From the last chapter (Chapter 4), it was learned that the transient GUS activity was reduced when some sequences were deleted from the 5' region of the promoter, which is similar to the observation that deletion of *Arabidopsis* PAL promoter from –832bp (relative to transcription start site +1) to –540bp and –290bp led to dramatic reduction of GUS activity in transgenic plants (Ohl *et al.*, 1990). At the transgenic callus stage, PAL400 showed similar GUS expression to PAL840 but PAL260 and PAL200 did not show any GUS expression in the limited number of lines analysed.

In transgenic rice plants, the expression patterns of GUS driven by PAL840 and PAL400 were the same, indicating that about half of the promoter sequence at the 5' end of PAL840 was not important for tissue-specific expression. This is consistent with the results from the deletion analysis of other PAL promoters in other plants. When bean gPAL2 was deleted to –420, -342, -283 and –254 (relative to transcription start site +1), the GUS

expression patterns driven by these deleted promoters showed the same patterns as those by –1190 promoter (Hatton et al., 1995). Similarly, gPAL3 deletion to –447bp showed GUS expression pattern similar to 3960bp promoter (Shufflebottom *et al.*, 1993).

The effect of the promoter deletions on the promoter activity would be much clearer when transgenic cassava plants are obtained from constructs PAL260GUS and PAL200GUS and fully detailed analysis of GUS activity is carried out.

5.4.2 PAL expression patterns in cassava and other plant species

It should be pointed out that one of the drawbacks in this research was that transformation of cassava with 35S-GUS was not carried out successfully, which makes it difficult to distinguish whether the observed expression patterns driven by the PAL840 represent the activity of the promoter or they are shared by other promoters or just a reflection of metabolic activity and cellular differentiation. However, by comparing the GUS expression patterns driven by cassava PAL840 with the expression data obtained by other researchers of 35S-GUS in cassava, and referring to GUS expression patterns driven by PAL promoters from other plant systems, the interpretation of the expression data of PAL840GUS may be more effective.

Although several laboratories have used 35S-GUS to optimise cassava transformation systems, they were interested only in examining the transgenic lines for GUS positive staining not the detailed expression patterns, therefore, the expression data were quite limited (Schöpke *et al.*, 1996; Li *et al.*, 1996; Zhang *et al.*, 2000 and 2001). A few weeks after transformation, dark blue (GUS staining) cells were visible in both 35S-GUS (Schöpke *et al.*, 1996; Li et al., 1996) and PAL840GUS (Figure 5.1a). At cotyledon stage, 35S-GUS showed strong GUS expression (Schöpke *et al.*, 1996; Li *et al.*, 1996), whereas most of transgenic lines containing PAL840GUS showed down-regulated GUS expression compared to callus stage and GUS expression was limited to the vascular tissue of the cotyledons (Figure 5.1c). In regenerated plantlets, PAL840GUS also displayed a different

pattern from 35S-GUS. GUS expression was only detected in the veins of the leaves in PAL840GUS plants (Figure 5.2 and Figure 5.3), whereas it was_expressed all over the leaves in 35S-GUS plants (Schöpke *et al.*, 1996). Most recently, it was learnt that GUS was expressed constitutively, in all the cell types, in transgenic cassava plants containing 35S-GUS (Taylor N, personal communication), in contrast, strong GUS expression driven by PAL840 was observed mainly in xylem parenchyma in leaves, stems, roots and tubers, and in cork cambium in the both shoot and root systems.

In a number of plants, fusion of the GUS reporter gene to PAL, cinnamate-4-hydroxylase, 4CL and cinnamyl alcohol dehydrogenase (CAD) genes has shown that the sites of promoter activity are generally consistent with expected sites of phenylpropanoid product accumulation. For example, in the cases of the bean PAL2 and poplar PAL2 promoter and GUS gene fusions, GUS was expressed in the xylem, a site of lignin deposition, and expressed in flower petals, a site of anthocyanin accumulation (Gray-Mitsumune, 1999). Most PAL promoters analysed by GUS fusion have been shown directing GUS expression in vascular tissues of various organs such as leaf, petioles, stem and root, where lignin is accumulated. Strong GUS activity driven by a rice PAL promoter ZB8 was located in vascular and epidermal tissues in transgenic rice, which was correlated with the presence of lignin or some other polymeric phenolic compounds identified by phloroglucinol staining in these tissues (Zhu *et al.*, 1995). In the present study, GUS was highly consistently expressed in the xylem of transgenic cassava containing PAL840GUS/MON gene fusion. It was also expressed in the cork cambium of the stem and tuber. Toludine blue staining showed in xylem and bark cambium tissue the presence of lignin or lignin-like phenolic compounds, where GUS activities were localized (see Figure 5.17 and Figure 5.18). The dramatic increase in GUS activity of hardening petiole and lignifying stems compared to those in young petioles and stems also demonstrated that the activity of MePAL2 promoter was in line with increased lignification of these tissues (Figure 5.4).

The difference between constitutive GUS expression driven by 35S promoter and the GUS expression pattern driven by PAL840, and the similar expression patterns between cassava

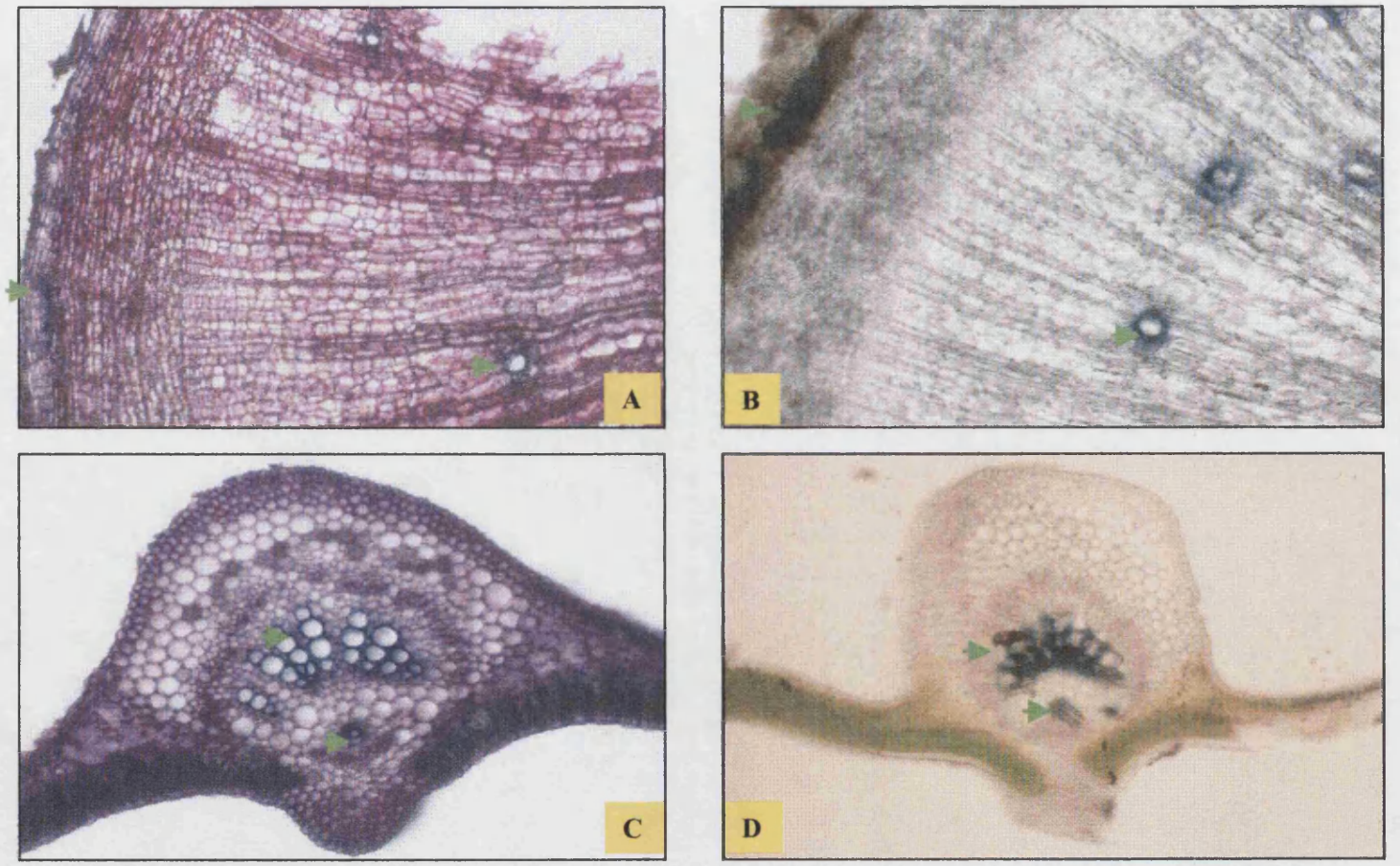


Figure 5.17 Comparison between GUS expression pattern and location of lignin in tuber (A & B) and leaf (C & D).

Cross-sections were from different tissues of 6 month old soil-grown transgenic cassava (PAL840GUS/MON). A & C section were stained for lignin using toludine blue, B & D for GUS activity. The location of GUS activity and ligninfied tissue were indicated with green arrows.

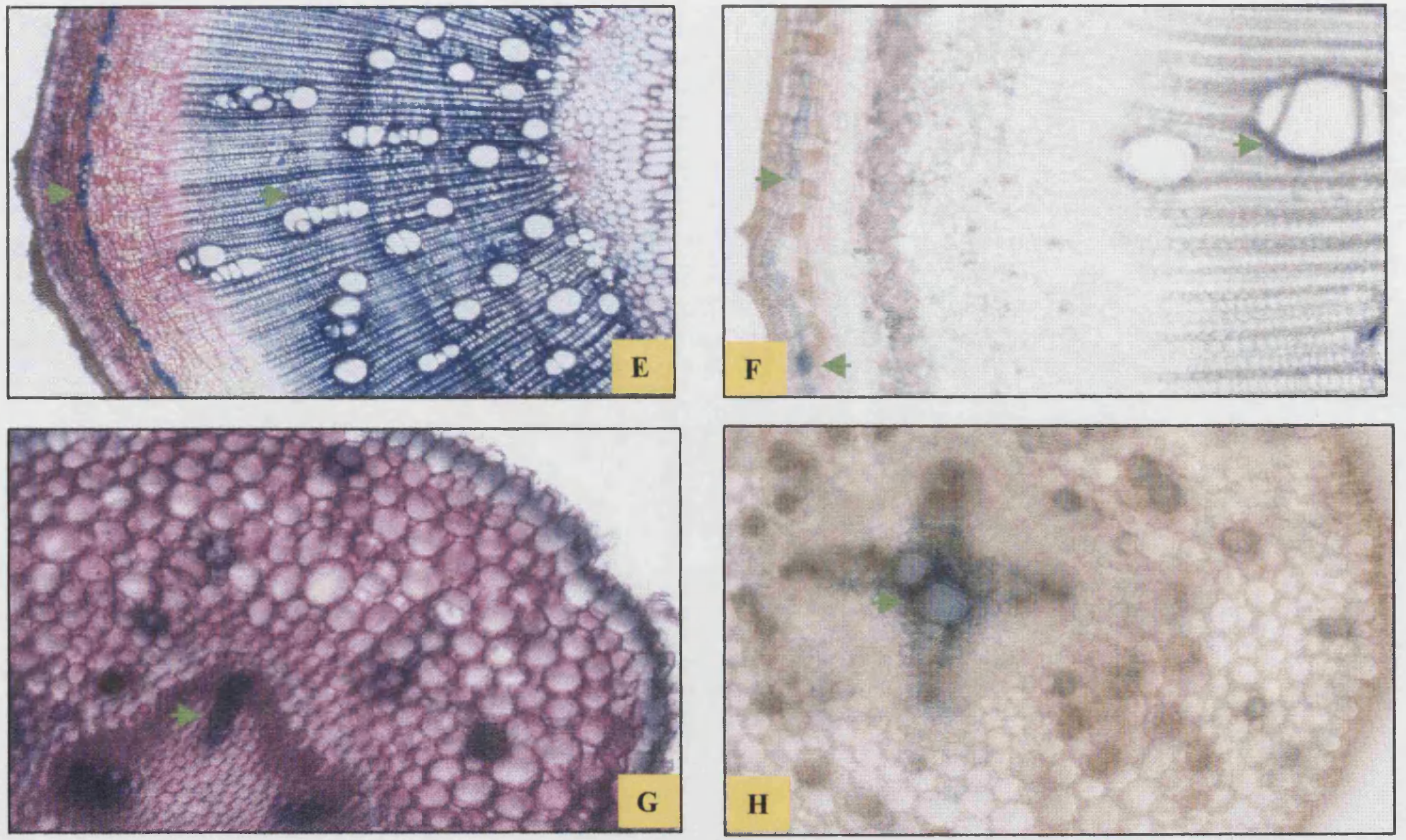


Figure. 5.18 Comparison between GUS expression pattern and location of lignin in stem (E & F) and vegetative root (G & H).

Cross-sections were from different tissues of 6 month old soil-grown transgenic cassava (PAL840GUS/MON). E & G section were stained for lignin using toludine blue, F & H for GUS activity. The location of GUS activity and ligninfied tissue were indicated with green arrows.

PAL840GUS and PAL-GUS from other plant systems, together with the evidence that strong GUS expression was localised to the tissue where there was lignification, suggest that the GUS expression driven by the PAL840 promoter is tissue specific.

The similar expression patterns among PAL promoters in plants may lie on similar *cis*-elements or domains such as P-A-L boxes (section 4.2) in the promoters. In order to understand whether there is any similarity in sequences apart from known *cis*-elements between promoters, PAL promoter sequences of poplar and cassava were aligned (Figure 5.19). The alignment showed that there is 57.9% similarity over 700 bp compared. Between these two promoters 11 motifs over 8 nucleotide long have identical sequences. These regions, together with putative *cis*-element boxes, may play a role in the regulation of promoters.

```
poplar   ACATGTAATGTAAGAATTCAGCAAACC---AATTATATAGGATAATTTTA--------TT
cassava  -CAATTAATATTTTAATTACAAAAATATTTAATTATATAATATAAAATTAAAATCTCATT

poplar   AGTAAGATTTATAATTTATTTTCTAAAAATTACTAGTTTAAATTTTTTTAAA------AT
cassava  GTTTAAATTTTTTATT-ATATGTTAAAAATTACATATTTTTATTTAATAAAAGGGTTTAT

poplar   TATTAAAAATTT----ATATAATTAT-TAATTTCAACATTCA-TAAAATTAGTTGAAATT
cassava  ATTTATATATTTTTTTTATAAAATTATGTCATATTTATACTAAATTAATTTATTAAAAATT

poplar   CATATAAATTCGCTTAAGAATTTATATTAATAATAATAAAAAAGATAGGCAATTTTGCGA
cassava  AAAATACTTTTAATGAATA-TTTATTTTATTTCTTCAAAAAAAAAGTTTCAATTTTACTT

poplar   -AGGTAAAAGCA----TGATTATTAGGTGGATCCA-TGTGTGTGTGTGTGT-GTGTGTGA
cassava  TAAAAAAAATCAAAATTTATTTTGAAATTTACTTACTGATTAGGTGGCTGTAGCGTTGGA

                                                      AC-I
poplar   -----AAGAAGGGCAAGACGAAATTTC----CCATTT-CTCACCAACCACAACTCGCCAC
cassava  GACCCAATCAATGGAA-ACGAA-TTTCTTGGCCAGTTGCCCACCAACCACAACCT--CAC

poplar   CATGCATCACACTACCGCGATAGTCAAATTTACCCTTCTACGCCAATCGCCAATATGGAT
cassava  CATGCACCACC--ACCACG--AGTCAAATTTACCCTTCTCTCCTAATCATCACTCT----
```

Figure 5.19 Alignment of MePAL2 promoter and poplar PAL1 promoter sequences.

Sequences of poplar and MePAL2 promoters 705 bp upstream of ATG translation initiation codon were compared. Putative *cis*-elememts or motifs in MePAL2 are in bold and underlined, and those in poplar PAL1 (AF038863) promoter are in bold. Identical nucleotides between these two promoters are linked by ':'. Identical motifs over 8 bp between two promoters are underlined.

A number of promoters from dicots have been shown to have similarly activity when they were transferred into different dicot species as they had in their species of origin. For example, PSPAL1 from pea, PAL1 and PAL2 from poplar and PAL2 and PAL3 from French bean promoter was active in tobacco (Gray-Mitsumune, *et al.*, 1999; Shufflebottom, *et al.*, 1993 and Kawamata, *et al.*, 1997). The monocot rice PAL, ZB8 promoter was shown to drive GUS expression in dicot tobacco (Zhu *et al.*, 1995). The heterologous activities of PAL promoters may reflect the conservation of some important domains such as P-A-L boxes in the promoters.

In rice, 35S-GUS showed constitutive expression (Battraw and Hall, 1990). GUS expression was detected in leaves (outer bundle sheath cells, thicker-walled inner bundle sheath cells and mesophyll cells, the xylem parenchyma cells adjacent to the metaxylem trachery elements, small parenchymatous cells of the phloem, leaf epidermis, trichomes and bulliform cells) and roots (vascular cylinder, thin-walled cells of the cortex and root tips). A rice *act1* promoter, more constitutive than 35S promoter, drove GUS expression in almost all cell types (Zhang *et al.*, 1991). In contrast, PAL840GUS in transgenic rice was expressed in outer bundle sheath, bulliform cells, stomata and trichomes of leaves, but not expressed in phloem cells, mesophyll cells, inner bundle sheath cells and epidermis. In root, PAL840GUS was expressed in root tips and vascular cylinder but not in cortex. This, to some extent, suggests that the MePAL2 promoter has its specific activity.

Rice PAL, ZB8 promoter showed similar tissue specificity. It was active in root apical tips, lateral root initiation sites, vascular and epidermal tissues of stems and roots (Zhu *et al.*, 1995). In order to understand whether the similar expression patterns of the cassava PAL840GUS and ZB8-GUS in rice are due to possible similarity between their promoter sequences, comparison between these sequences were carried out. Comparison of the promoters from MePAL2 and rice ZB8 PAL over 900 bp showed that they shared a similarity of 48.3%, which is lower than that between poplar PAL1 and cassava MePAL2. There are 5 regions (over 6 bp) showing identical sequence between these two promoters and they share the same putative TATA box (Figure 5.20). The P-A-L boxes in cassava MePAL2 promoter were not found in the rice PAL promoter. Interestingly the AC- rich regions (covering 120 bp from TATA upstream) of these two promoters have similar AC ratio 62 AC/95 bp (65.3%) for rice and 65 AC/102 bp (63.7%) for cassava. It has been shown that AC-rich motif of bean PAL promoter is implicated in xylem expression and a transcription factor binding the AC-rich motif has been isolated (Sequin *et al.*, 1997). Therefore, the active expression of cassava PAL promoter in rice may indicate that possible transcription factors of PAL promoters from different plant species may be conserved.

```
rice      T----CGCGTCTAACTGACCTACGTATGGAAATGAAGCGGAGACTTGTTCGGCGACATTG
cassava   TTAAACTAAGCTTATTGTGTTATTTATTTTAATAATAATAATAATTTTTACCCAAAAAA-

rice      ACTGGAACGTGGTCTTATGTGTGCGCACAATCATATTCAGGTCTCGCTGATATGTTCAAA
cassava   ACAACAACATAT--TTAACTCTTACCA-ATACATTCTCACTTTATATTTATAT-TTATAA

rice      CTAGTAAATACCTTTTGAATATCTTCGCCGGCCCGGCCCCGGTGAACGAAATGCATTCTG
cassava   CAA---AATACTATAAGAA-ATTTTA------------------ATAAAAATAAAAGATA

rice      AA--TAGGGCATGCACCGAGCAA---CGAATACTGAATTGCAAATTGTTTCAGCTAGCTT
cassava   AAACTATAGACTAAAATTAGCAAATTCGGTTAATTAGCTTCAATTTTTT--AGTTA--TG

rice      TACAAAATTTGACCGGTTGTTCTATCAAAAAAAAAAATTGAACGGTTATTTCCTGAGAAG
cassava   TACAACACACAATTAATATTTTAATTACAAAAATA--TTTAA---TTATATAATATAAAA

rice      TTAGACTCATCAAACATTA---TTTTTAGGGGCACTAAAGGCAGAAGCCATGTCTCGTTC
cassava   TTAAAATC-TCATTGTTTAAATTTTTTATTATATGTTAA---AAATTACATATTT---TT

rice      GTTTGATAAAAGAAATGATAC----ACATTGACCAAAAAAACCGGATCACA------CCA
cassava   ATTTAATAAAAGGGTTTATATTTATATATTTTTTTATAAAATTATGTCATATTTATACTA

rice      ACTTCCATCCATAAAAAAAAGGAGTGCTACCAACCGAGCAACGGGT--AGTCAGTCATAG
cassava   AATTA-ATTTATTAAAAATTAAAATACTTTTAAT-GAATATTTATTTTATTTCTTCAAAA

rice      CAATTCTCTCCAGGACGAGGAGTACCAAAAGCAGCAGCACTGTGACGACTGACGAAGTGA
cassava   AAAAAGTTTCAATTTT-ACTTTAAAAAAAATCAAAATTTATTTTGAAATTTACTTACTGA

rice      CGAGCACGAGCCAGTAACG--GGAGACCGGGCAGCCGCAGACGCAGCAAGTCCGTCGCGC
cassava   TTAG---GTGGCTGTAGCGTTGGAGACCCAATCAATGGAAACG----AATTTCTTGGCCA

rice      GC--CCCACCACCAACACCCCGTGCCGCGCAGGCCCACCAACCACCACAC--AGCCACAG
cassava   GTTGCCCACCAACCACAACCTCA-CCATGCACCACCACCACGAGTCAAATTTACCCTTCT

rice      CGCCG------CACACGAA---ACTTCCACACCGCAGAG---CCCAAA--AACCTCCCGT
cassava   CTCCTAATCATCACTCTCATGCAATCCCAACCCTTGGATTTTCCCAAATCAATGGCCATT
```

```
rice      TTTC-TCTCCTACTCCGCGACG--CTGACCGTGTGCTCTCGCCTTGGATCACCAGACCAT
           ::  : : : ::    : ::   ::  :: : ::: ::: : :  : : :::   ::::
cassava   ATTAATTTTCAACCAACCCACCTTCTCTCCCTCTGCCCTC-CTTATGCTTACCTACCCAT
                  P-box                      A-box        L-box
                                                          TATA box
rice      CACCAACCCCCCTC-------CGTCATCCGTCCTGCAGGCCACCCACCCGCCTATTTAA
           ::  ::     :        : ::: :  : :: :   :  ::::::  : :::::::
cassava   TACACACTTATATGGAAGTCTCTTCACCTCTACTTCTCTCTCCCCACCTTC-TATTTAAA
                                                          TATA box

rice      TATTTAAGCCGCGCCTCCCCCTCCATTCCCCTCCAAGAAGAGCCGCACCATCCAGTGCAG
          ::::::: :  : :  :  : :::  : : :  :: :::          ::::: : :
cassava   TATTTAAACTCCACTCCTTCATCCTCTGCTCCTCAGGAA----------ATCCATTTCC-

rice      TAGTACACTAGCTCTTCTTCACAACAGCTAATCGAGTAG--CTAGAAC--CATTATATAC
             :::   :: : ::: :: :  ::: :  :    :    ::  :::  ::   : : :
cassava   ---TACCAAAGAT-TTCCTCTCT-CAGATCCTTTGTTTCTCCTTCAACTTCAAGTTCTTC

rice      TCTTCTCTCGACGCTTTTTG-TG-CTAGGTTAACCGATCCATCTTCTGGTACTGAAATG
           :::: :: ::    ::::: :: :: :::      ::     ::::  ::   :::::
cassava   ACTTC-CTTGAGTTGTTTTGCTGTCTGGGT------AT-----TTCT--TAA--AAATG
```

Figure 5.20 Comparison of rice and cassava PAL promoter sequences.

The promoter sequences 780bp upstream of the putative TATA box and the sequences between TATA box and ATG translation initiation site were compared between the two promoters. Putative *cis*-elememts or motifs in MePAL2 are in bold, and those in rice Z8 PAL promoter are in bold. Identical nucleotides between these two promoters are linked by ':'. Identical motifs over 6 bp between two promoters are underlined.

From these two alignments of promoter sequences, between cassava and poplar or rice, it is also noted that the second putative TATA and nearby sequences in the cassava MePAL2 promoter align well with poplar and rice sequences in the regions of TATA box (5.19, 5.20), which supports the suggestion made by TATA mutation and transient assays that the second TATA box was the functional TATA box.

5.4.3 MePAL2 expression in response to wounding

A number of PAL genes in other plant systems were regulated by stresses such as wounding, pathogen infection and UV light irradiation (Dixon and Paiva, 1995; Liang *et al.,* 1989; Lois and Hahlbrock, 1992; Reinold and Hahlbrock, 1996; Subramaniam *et al.,* 1993; Wu and Hahlbrock, 1992). It has been shown that wounding can enhance PAL expression in many plants such as rice ZB8, bean PAL2 and PAL3 (Zhu *et al.,* 1995; Cramer *et al.,* 1989). PPD is a special wound response with little or no wound healing

(Beeching *et al.,* 1996). Therefore, GUS expression in transgenic cassava tubers containing PAL840GUS and undergoing PPD was analysed. The MePAL2 promoter appeared to be regulated by PPD. Interestingly, the location of GUS expression in tubers was where visible symptoms of PPD developed. The patterns of GUS staining in cross-sections of tubers during different stages of PPD showed no discernible difference. However, GUS activities measured by the MUG assay showed an increase within 24 hours post-harvest, which decreased slightly by 48 hours and then remained at the same level until 72 hours. This may indicate that the activity of the MePAL2 promoter is affected by PPD. Furthermore, pruning two weeks before harvest delayed the appearance of peak GUS activity during PPD, which again suggested the regulation of this promoter by PPD, as pruning can delay PPD (Tanaka *et al.,* 1984). Tylose formation has been associated with wound response and defense response, and tyloses were thought to be involved in the prevention of water loss (Fahn, 1982). In cassava stems and roots, we observed tyloses, this is the first report of tylose formation in these organs though it has been reported in the cassava tuber (Rickard, 1983). GUS was strongly expressed in the tyloses of stems, roots and tubers, also indicating the possible regulation of gMePAL2 promoter by wounding. However, it cannot be concluded that the response of MePAL2 to wounding is specific without comparing to the GUS expression driven by constitutive promoters upon wounding, as there are reports that constitutive promoters can direct non-specific increase of the expression of reporter genes upon stresses such as increased GUS activities driven by 35S promoter in transgenic tomato leaves upon wounding (Blume and Grierson, 1997) and maize ubiquitin promoter in transgenic rice upon heatshock (Cornejo *et al.*, 1993). To test whether the PAL840GUS is regulated by wounding, jasmonic acid (JA) may be applied. JA has been shown to be one of the main components in signalling pathway of wound response (León *et al.*, 2001). Exogenous application of JA or its methyl ester induced wound responses such as the expression of wound—responsive genes including proteinase inhibitor gene *Pin2* in tomato and vegetative storage protein gene *Vsp* in soybean (Farmer *et al.*, 1992; Mason and Mullet, 1990). Blocking the synthesis of JA by salicylic acid or aspirin inhibited the wound-induced accumulation of PR genes (Tomoya *et al.*, 1998). Lanolin containing JA or linolenic acid, a precursor of JA biosynthesis, may

be applied to transgenic cassava tubers containing PAL840GUS without detaching them from the plants to avoid artificial mechanical wounding, GUS activity in the tuber of the treated plants can be compared to that in the tuber that is treated with water to find out whether the PAL840 promoter is responsive to the JA treatment. A parallel control experiment should be carried out in transgenic cassava containing 35S-GUS (available from ILTAB, Danford Center, St. Louise, USA), which may help to clarify whether the response of PAL840 is wounding-specific or not. However, it would be difficult to avoid the possible damage to the tuber and plants when tubers are treated in the way mentioned above, which may cause stresses to the plant and the tubers. An alternative approach or supplement experiment would be to apply JA or its methyl ester to the PAL840GUS and 35S-GUS transgenic plantlets in Magenta box and analyse the GUS expression of leaves after the treatments. This should be preceded by the analysis of GUS activity before and after mechanical wounding of the leaves in PAL840GUS and 35S-GUS. Through these experiments, a convincing conclusion may be drawn about the specificity of the cassava MePAL840 promoter in wounding response.

5.4.4 Variation of GUS expression among different transgenic lines of the same constructs

Variation of GUS expression was observed among different PAL840 lines of transgenic cassava and rice. There several factors possibly related to this phenomenon, which include copy number and position of the transgene in the genome and gene silencing.

In order to investigate whether the copy number of transgene were the cause for the variation in GUS activities among different lines, Southern hybridisation of the genomic DNA from randomly chosen cassava and rice transgenic lines was carried out. Preliminaries Southern analysis of transgenic cassava suggested that the transgenic cassava lines may be different in the copy number of the transgene, although the quality of the Southern was poor (Figure 5.21).

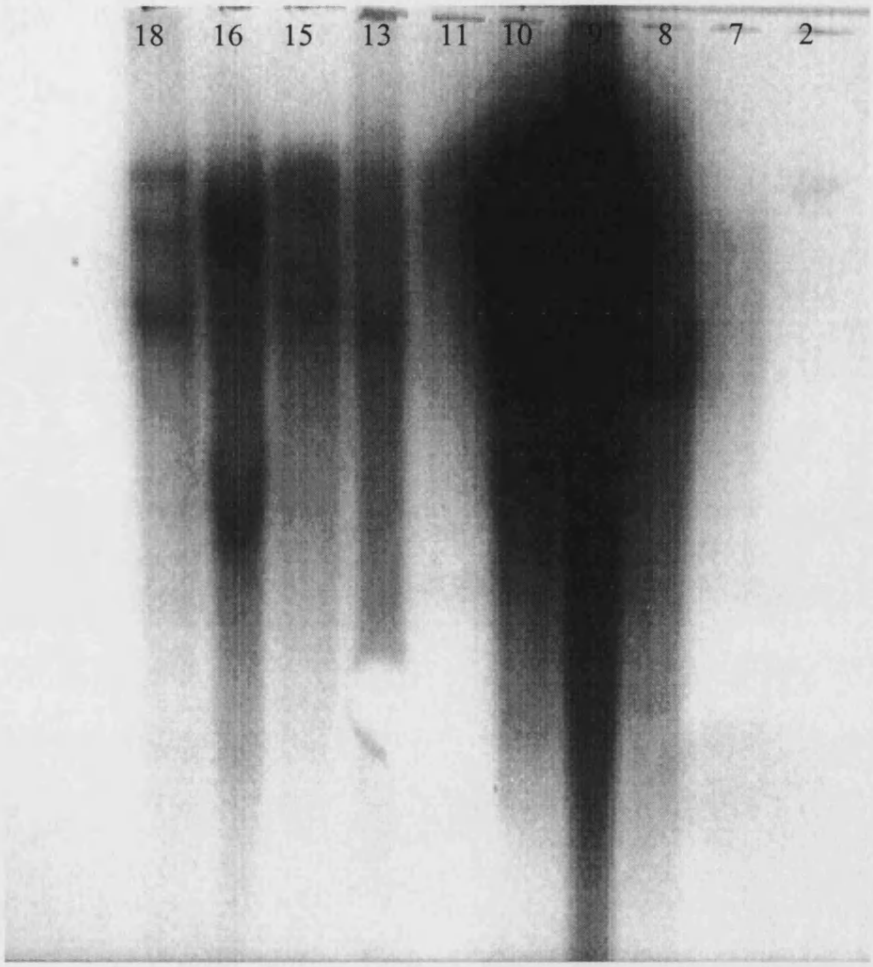


Figure 5.21. Southern analysis of 840PAL-GUS transgenic cassava lines.

Genomic DNA was extracted from young leaves of 840PAL-GUS transgenic cassava lines 2, 7-11, 13, 15, 16 and 18. The genomic DNA was digested with *Hin*dIII, run in 1 x TBE 0.8% agarose gel and subsequently blotted against Hybond-N^+. Hybridization was carried out at 60°C overnight as described in section 2.3.5 (chapter 2) and the probe made from *udiA* gene. After hybridization, the membranes were washed finally stringency (60°C, 1 x SSC ,0.1% SDS, 2 x 10 min). The membrane was exposed to film between two intensifying screens overnight.

Among transgenic rice lines, it appeared that lines showing several bands in Southern such as Line 3, 6 and 7 showed no or low GUS expression, whereas Line 2 that showed one band in Southern showed high GUS activity. However, Line 4 that showed multiple bands also displayed high GUS activity (Table 5.3; Figure 5.22).

Line	1	2	3	4	5	6	7	8	9	10
GUS activity#	+	++	-	+++	-	+	-	+	-	-
Transgene copy no.	6	1	>6	5	3	>6	>6	3	-	-

Table 5.3 GUS activities and loci number of transgenes in transgenic rice lines Note: GUS activity was visually measured according to the intensity of GUS blue after staining, '+' indicates blue staining was visible; '++' indicates stronger staining and '+++' for strongest staining. '-' no visible staining for GUS staining or no transgene was detected in Southern hybridisation.

The lower GUS expression in some lines containing multiple copies of the transgene than Line 2 containing 1 copy of the transgene, may possibly be due to sense gene silencing. There are numerous examples of sense gene silencing in transgenic plants (Napoli *et al,* 1990; van der Krol *et al.*, 1990; Smith *et al.*, 1990; Baulcombe, 1996 and 1999) and it is much more frequently observed in transgenic lines with multiple insertion of the transgenes (Flavell, 1994). In rice, the aleurone-specific Ltp2-GUS transgene was silenced. Cloning-out of the transgene demonstrated that two copies of transgene were inserted into one locus, one intact Ltp2-GUS and the other was a partial sequence of the transgene in the antisense direction, which was suggested to be the cause of the silencing (Morino *et al.,* 1999).

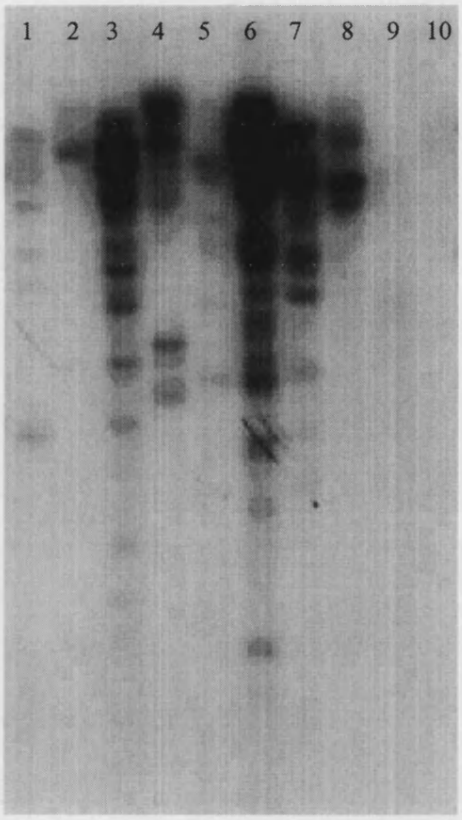


Figure 5. 22. Southern analysis of 840PAL-GUS transgenic rice lines.

Genomic DNA were extracted from young leaves of 840PAL-GUS transgenic lines 1 to 10. About 10 μg of the genomic DNA was digested with *Hin*dIII, run in 1 x TBE 0.8% agarose gel and subsequently blotted against Hybond-N^+. Hybridization was carried out at 60°C overnight as described in section 2.3.5 (chapter 2) and the probe made from *udiA* gene. After hybridization, the membranes were washed finally stringency (60°C, 1 x SSC ,0.1% SDS, 2 x 10 min). The membrane was exposed to film between two intensifying screens overnight.

5.5 Summary

The promoter region of MePAL2 was analysed and a number of putative *cis*-elements such as P-A-L box and motifs with specific features such as AT rich regions were identified. 840 bp of the MePAL2 promoter, containing P, A and L motifs was fused to the reporter gene β-glucuronidase (GUS) (PAL840GUS) and transformed into cassava, as were deletion mutants (PAL400GUS, PAL260GUS and PAL200GUS) of the promoter in which successive motifs had been removed. The PAL840GUS was studied in detail. It showed expression in the xylem parenchyma cells of the vascular tissue of all tissues examined and, in some cases, in the cells of the epidermis and bark cambium. In the storage root, the MePAL2 promoter seemed to be increased during physiological deterioration, with expression confined to those vascular tissues in which the visual discoloration observed during PPD occurs. The vascular expression patterns of cassava MePLA2 promoter are similar to those of other plant PAL promoters, which may lie on the conservation of motifs such as the P-A-L boxes and possibly transcription factors. It should be pointed out that the properties of the MePAL2 promoter may not be limited in these aspects observed with PAL840GUS, as a number of putative *cis*-elements such as two G-boxes (wounding responsive) and an ethylene inducible motif were located upstream of the PAL840 (Figure 4.1). Deletion analysis suggested that PAL400GUS displayed the same expression pattern as PAL840GUS. Further deletion of the promoter led to the lost of its function in transgenic callus, though the deletions were transiently functional in suspension cells. The deletion study confirmed that P-A-L box or AC rich motifs were critical for the expression of the PAL promoter.

Additionally, the PAL promoter GUS constructs were transformed into rice. Analysis of those plants that had received the PAL840 bp promoter indicated that its expression during development was similar to that observed in cassava.

CHAPTER SIX: ISOLATION AND CHARACTERISATION OF AN ACC OXIDASE GENE FROM CASSAVA

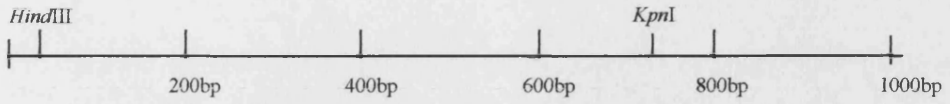
6.1 Introduction

Ethylene is involved in a wide range of physiological processes in plants including germination, abscission, senescence, flowering and fruit ripening. Ethylene production is also induced by stresses such as wounding, drought, chilling and pathogen attack (Miki *et al.*, 1995; Lasserre *et al.*, 1996; Pogson *et al.*, 1995; Gomez-Cadenas *et al.*, 1998; and Abeles *et al.*, 1992). During cassava post-harvest physiological deterioration (PPD), ethylene production in the roots was found to increase after a lag of 6 to 16 hours after harvesting (Plumbley *et al.*, 1981; Hirose, *et al.*, 1984). 1-aminocyclopropane-1-carboxylate (ACC oxidase or ACO) catalyses the last step of ethylene biosynthesis in plants. A number of ACC oxidase genes have been cloned, mainly from ripening fruits or flowers and some from vegetative tissues of different plant systems (Lasserre *et al.*, 1996). In most plant systems studied, ACC oxidase is encoded by a multigene family. An ACC oxidase cDNA clone, designated as MeACO1, has been isolated and characterized from a PPD-related cDNA library (Han, 2000). MeACO1 was 1006 bp in length and encodes 324 amino acids. The deduced amino acid sequence showed up to 85% similarity to tomato pTOM13 (ACO1) and tobacco NtACO. This chapter reports the estimation of the family size of the ACC oxidase gene in cassava, and the isolation and characterization of an ACC oxidase genomic clone.

6.2 ACC oxidase gene organization in cassava

In order to estimate the number of ACC oxidase genes present in cassava, Southern analysis of cassava genomic organization of ACC oxidase gene was carried out.

Ten micrograms of genomic DNA was digested with *Eco*RI, or *Xba*I, or *Hin*dIII, or *Sac*I. These restriction enzymes were chosen as *Eco*RI, *Xba*I and *Sac*I do not cut MeACO1, whereas there was a *Hin*dIII restriction site at the very beginning (39 bp) of the MeACO1 (Figure 6.1). The digested DNA was run in agarose gel and transferred to Hybond N+ membrane. Then the membrane was probed with cDNA MeACO1 following the protocol described in section 2.3.6. From the hybridization result (Figure 6.1), it can been seen that there were at least four bands when digested cassava genomic DNA was hybridized with the cDNA MeACO1 probe. There are four positive bands in the *Eco*RI digestion lane. One of these bands, about 8 kb in size, with a stronger signal than the others, may be from the genomic DNA of the cDNA MeACO1, whereas the weak signals may be from other ACO genes with high similarity to MeACO1. *Xba*I digestion resulted in five positive bands and one of them showed a much stronger signal. *Sac*I digestion led to 4 positive bands with large sizes, from 9 to 23 kb. *Hin*dIII digestion showed three weak bands and two stronger bands. Though there are no restriction sites for those enzymes (except *Hin*dIII) in the cDNA probe, it might be that there are these restriction sites in introns of the genes. It may be suggested that there are probably at least ACC oxidase genes in cassava.



A: The map of cDNA MeACO1 (1006bp). *EcoR*I *Xba*I and *Sac*I did not have restriction site in the cDNA MeACO1

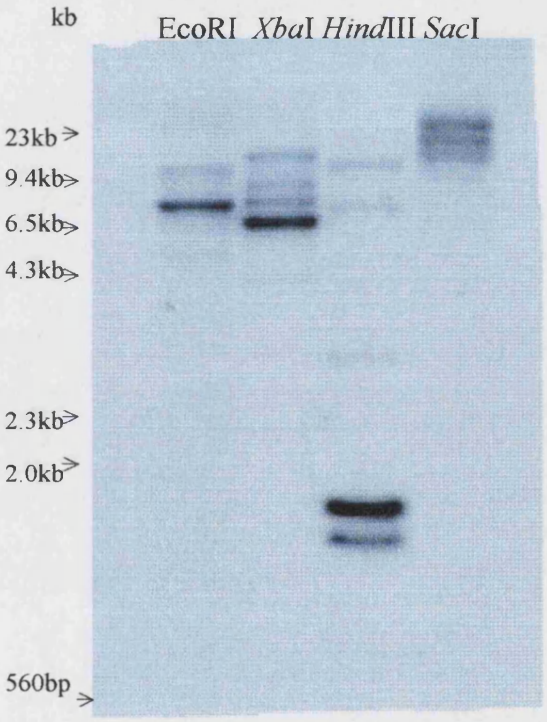


Figure 6.1 Genomic Southern hybridisation using a probe from the full length MeACO1 cDNA

Genomic DNA was restriction-digested with *Eco*RI, *Xba*I *Hin*dIII and *Sac*I respectively and Southern-blotted to HybondN+ membrane. Hybridization was carried out at 60°C overnight. Then the membrane was washed at low stringency (60°C / 1xSSC / 0.1%SDS / 2x10min) and exposed to Kodak film overnight. The restriction map of MeACO1 cDNA is shown in the top for reference.

6.3 Isolation of an ACC oxidase gene from a cassava genomic library

About $2.3x10^6$ plaques of the cassava genomic library were screened with the cDNA MeACO1 probe (see section 2.4.2). Only one positive clone was retrieved from the first genomic library screening. After the second and third screening, separate positive plaques were obtained. To confirm that the positive clone corresponded to ACC oxidase, PCR was carried out using the lysate of the plaque and primers cacof (AAG ATA TGG ACT GGG AGA GC) and cacor (CAT TGC CCA TCT TTC AGG AG) designed from cMEACO1 (for the position of the primers see figure 6.2A). The PCR product amplified from the positive clone, was about 500 bp, which was slightly bigger than the PCR product (300 bp) from cMeACO1 (Figure 6.2). If the genomic clone corresponded to cMeACO1, there must be one or more introns in this region. The PCR product from the genomic clone was purified and sequenced. The sequence was submitted to a Blast search, which showed that this clone had high homology to ACC oxidase genes from other plants and that these genes did contain an intron in this region. However, when compared with the cDNA clone, MeACO1, the similarity was 90% over a 300 bp length. Therefore, the isolated clone was an ACC oxidise clone but not the one corresponding to cDNA MeACO1. This genomic clone was designated MeACO2.

A: MeACO1 cDNA map and the positions of the primers used for PCR

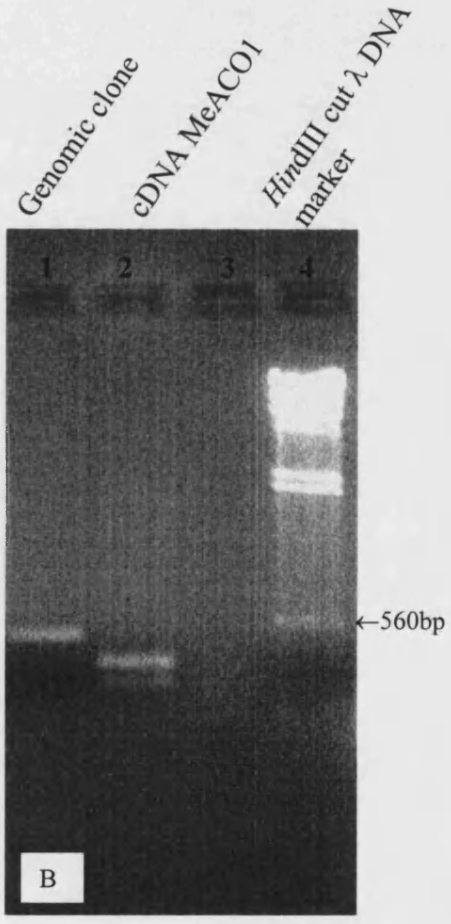


Fig 6.2 PCR test of the putative MeACO genomic clone.

A: cDNA MeACO1 map. Acof (AAGATATGGACTGGG AGAGC) and acor (CATTGCCCATCT TTC AGGAG) were two primers used for confirmation of putative ACC oxidase genomic clone.

B: PCR products from cMEACO1 (control) and genomic clone.

1 is putative MeACO genomic lambda clone as a template, and **2** cDNA MeACO1 as a positive control. PCR condition: 94 °C 3min 1 cycle, 35 cycles of 94 °C / 1min - 58 °C / 1min - 72 °C / 1min, 1 cycle of 94 °C / 1min - 58 °C / 1min - 72 °C / 5min.

In order to identify appropriate fragment(s) for sub-cloning and subsequent sequencing, the lambda DNA was extracted, purified and digested with *Sac*I, *Kpn*I, *Hin*dIII, *Pst*I, *Xba*I and *Bgl*II. The digested lambda DNA was analysed by Southern hybridization with the full length of cDNA MeACO1 as probe. It can be seen from the hybridization result (figure 6.3) that 6 kb and 2 kb fragments were released by *Hin*dIII and *Kpn*I digestion, respectively, these were of appropriate sizes for sub-cloning. Though *Sac*I, *Pst*I and *Xba*I digestion released single bands containing the whole ACC oxidase clone, the bands were above 10 kb, which would be difficult for sub-cloning. *Bgl*II released a small band less than 1 kb in single, but its signal was weaker than the 10 kb band, indicating that the small band may contain just a very small fraction of ACC oxidase DNA sequence. Therefore, the 6 kb fragment from *Hin*dIII digestion and 2 kb fragment from *Kpn*I digestion were purified and cloned into the pUC19 vector using the *Hin*dIII and *Kpn*I sites, and designated as phACO and pkACO respectively. The inserts of phACO and pkACO were sequenced using universal primers first, and then specific primers designed from interval sequences.

The individual sequences were edited to remove ambiguities and assembled to give the complete gene sequence (this was done by Novartis, San Diego, USA).

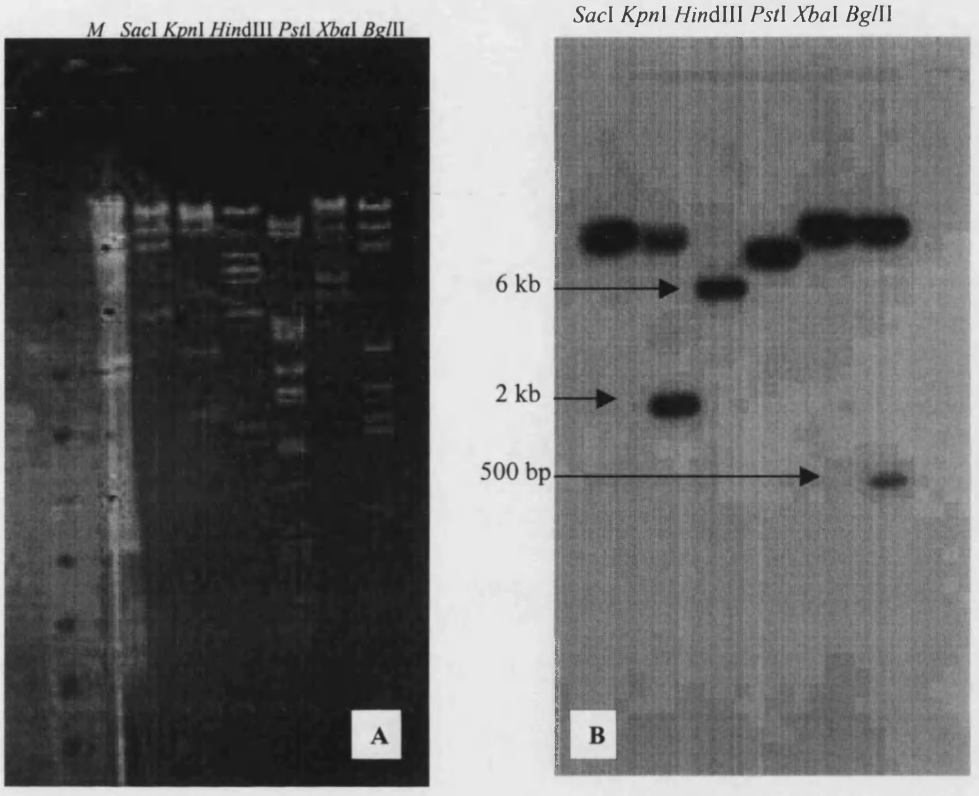


Figure 6.3 Southern hybridisation of restriction-digested genomic clone MeACO2 using MeACO1 cDNA probe.

A: The gel of MeACO2 clone digested by *Sac*I, *Kpn*I, *Hin*dIII, *Pst*I, *Xba*I and *Bgl*II respectively. M:*Hin*dIII cut λ DNA marker

B: The digested DNA was blotted to HybondN+ membrane and hybridised with MeACO1 cDNA probe. Hybridization was carried out at 60°C overnight. Then the membranes were washed at high stringency (60 °C 0.01 x SSC 0.1% SDS 2 x 10 min).

6.4 The Structure of MeACO2 gene

The determined sequence of the ACC oxidase gene, MeACO2, was 6349 bp, including 4928 bp of 5' UTR or 5' flanking region and four putative exons interrupted by three introns (Figure 6.4). The deduced amino acid sequence was 312 amino acids. The junction between intron and exon followed the /gt-----ag/ pattern, which is conserved in plant systems (Brown, 1986). Translation start site was determined using ORF-finder at http://www.ncbi.nlm.nih.gov/gorf/orfigure.cgi. Putative exons, introns and promoter regions were identified by Genefeature at http://vega.igh.cnrs.fr.bin.align-guess.cgi, and by comparing the gene with other ACC oxidase genes.

```
-4928  TGGAGTGCACGTGTATGATGGGTTCTAAAATGTTTTTTTTTTTTAAATTAATTTTATTAA  -4869
-4868  TCATATTTTAAAAAAATAAACTATCATCTAAAATATTGACTAAAAAAAATGTAGAATAATA  -4809
-4808  TCAAAGCATAATTGATACAAAATGAACTTAGACCTCAAATTAAAAGGTAAGATTACCATA  -4749
-4748  TTAAATAACGGTCCTAATTAATAAATAGAGATCAATTAAATCATTAGGCTAATTAGTTTC  -4689
-4688  TATTATTTTGAACAATCATACTGAATTCGAACAACTCATTTAACTTAAACAGACTCCATT  -4629
-4628  CCTTTTTTAACATTCATTTTAATATCAAACGGGTGACATAATCTTATTTAACCTTATTTT  -4569
-4568  TTAACCACACAACACAATAATCTCTTTGAAAATAATGCTGAATCTCTTCGAAAATAATGA  -4509
-4508  CTTCAGCTAATTGTGAATTGAAAATACCCAGTCGATGTCCTGTCAATCCCATTAAGTGCC  -4449
-4448  ATTCGAATTCCAAGCTACAGCACCAAAATTCAATTTCCCTTCTTGAGCAAAAACTCCAGC  -4389
-4388  ATCTAGATTATATTTTAATATGCCCAAGATAGGCTTTACCCAACTGGGATTGTTTGGTTG  -4329
-4328  TGGGAGGGCCTTCATTTCCTTTTTTCAGTCTGCCAAACTTGCTAAAGCTCTATCTACAAT  -4269
-4268  TAGAGAGGAACTAACTCCTTCCCATACTACAAACTTTTATTCTTCTAATTCTAAATATTC  -4209
-4208  CACATGATCATCTAAATTTTTGACCCGCATATCAATATCAAAATTAATATAGATAGATTA  -4149
-4148  ACGTACATCCTCAAGAGACTTAGTACTGAGTACCGATCGAGTTTTAATTGGTTAGTAACA  -4089
-4088  AAAAATTTTAAGTTTTTTCAATATTCCAATACTCTTATTGTATAGTTAATGATTTTTTAG  -4029
-4028  GTGTAGAATTAAAATTTTGAAATTGGGTTTGATTTTTATGCAAAATTTAGTCTCTTCTTG  -3969
-3968  GGGTCACTTTCTGTTCATCAGCATGTCTTCTATCTCATGTGGTTGGCATTGTGGATAAGC  -3909
-3908  ATAGGGCAATAAGAAAGCGAAGAATGAATGAGAAGTAAATTGGAGGAAGCATATGATTCT  -3849
-3848  TAGTTGTAAGTATGAAATGGACGTTTTCCTGTTACATGCAAGGTATAGCTTTTCTATACA  -3789
-3788  TCTCCTTTCCATTGAAAATGAAATAGCATTTTCAAAGTAGTAATACTATTTTAGTGGGGG  -3729
-3728  GGGGTTCCACATTTTGCCTCCAAATGTGACTTACAATGAAATTCTCTACTTTTGCCAATC  -3669
-3668  TCTATCTCTGTTCCATTTGAATGTGATGTTGAGTATAAAACCAATAAAATAGAGGTTCCA  -3609
-3608  TAACCTGAAAAACCCTAGTTGTTGTTCTCAATTATCAGTTAAAAGATTCCTCTCTTCCAT  -3549
-3548  TTCTTTTTTGCTATAGAAAGATACACTCGTGTATAATGTATACCCGGCTATTGAATATAA  -3489
-3488  TTAGTTTTAAGACCTATTATGATTTTAATTTGTTTGTAATTTGATGACCGCTGGATTCCT  -3429
```

-3428 CCACTCCAGTCAGGGGCCAAGTCCTCGATCGCGGCCCATCGCTCAGAAGCCCACACGTTC -3369

-3368 TCTGCACTAACAGGTCAGGCCCTTTCGAGCCTTGATAACGGGCCCCTCCGAATTTCTTAA -3309

-3308 CCAGGCCGGCCTCACCTTTACTGCACCTCGGGCCGATCTCCTTTGGGCCCAGCTTTGCTC -3249

-3248 CTATCTTCCGGCCTAGTCCAGAACCAGAAGGAAGACTCAATCATCTGCCTTGTGGACCCC -3189

-3188 TCTATACACGCACAGGGAGAATCAGAGGTCGTTACACATGGAGCAGAGATCTGATTCCCT -3129

-3128 TGTACGTCCGTATCAATGTAGCAGGGACAGATGGACCAACGATATACGTTCTGTTAGCAC -3069

-3068 GTCACTAGCAGACAAAAGGAAACATATAAAAGGAGAAATACTCTCATTTAGGTCTAAGTT -3009

-3008 TTTCATTATTCCAAAAACTTCTGTAAAACTCTATTTTTTGGATTTCAGATCATCATAATT -2949

-2948 TATTTATTTATTGAATGTTTATATATGCAGGCCGATACATGTAAAAATTTCCTTTATCCA -2889

-2888 CACGAACCGAGTAATTTCTTTGCTTTTAATTTCTTTGCTAGTTAGATAAATTACACGTGG -2829

-2828 TCAAATGGTTCATCAAAATAGAACCAATTTCTTCATTGGTTATTTTTGCTTTGAGGATAA -2769

-2768 CATTACATGCATTCCAAATAAAGTGATAAACCAATTTTTTGGAAGCTCATCAATGCTTTT -2709

-2708 CAATCTTTCCTCACCACCCACATGCAAAAAAAGGCAATACATGAAGTGTCCAAAAGCCAA -2649

-2648 GGCCAGCTTTTCTATTATTAGTTAGACTTAAATCATGTGGAATCTACAGTAACAAAGGCT -2589

-2588 TAATAATATGTTCATAATTGATAGATTTGGTATTCTATTACTGGACCAAACAAGTCGGAA -2529

-2528 GAAATACAGAGATGCTGGCTTTCCTTTTATAGAAACAAAATACTATTGAATCATGAGCTG -2469

ATTTCAA motif

-2468 GTGGGATTTTATTTTATCAGATGGTTGTGATTATTCT**AATTCAAA**ACAGTATATAACTAA -2409

-2408 CCATTCACTCTTTAGTTGAGTGAATAATATTAAAAAGTTAGTCGAGTGATTTAAAATTCA -2349

-2348 ACTTTTCATGGTGATTTATTAATTAGTTAATTAATTGTGTGGTCTTTGAACATATAGAAT -2289

-2288 GTATATGATTGAAGAAAATAAACATAGCATGTGCAGAAGGGGATAGAAATTTCCAGAAAT -2229

-2228 AAAAAGTGGTCCAAAGTCTCCAGCACAGACGGTGGTGACCTAAAAATTAGCAAAATAGAT -2169

ATTTCAA motif

-2168 GCAGTCAGCATGTAGAAAATTATAGGAGATTGCATCATCAT**ATTTCAA**GGAATTAATCAA -2109

-2108 GAAAAAACCAAGTCCACTTGATTGCTCACATGGTGATGAGCTCATTGTCATACTCATAAT -2049

-2048 TTTTTTTTTATTTATTAAATCTACGAGAAAAATAAGGAATTCAGAGAACCTGGAACTTCT -1989

-1988 TCCTGGACCAATTACACACATGTGTTGTCACAAAGTAGGAGGAGAAATCACAGTTCTTTT -1929

-1928 GACTTTCTTTTTCCTTTTCTTTTTTTTTGAAAAAATAAAATATGATACTATTTTCATGTA -1869

-1868 TTTTTATAAAACATTTACAAGTGTTAGATTTGTAAAACAACAACTCAGTGCCTAACTGCC -1809

-1808 GAATGAATGGTCATAATCATGAATTGAAAATTAAAATGGATTATTATATTTTCTAATTCC -1749

-1748 GTAAAATTTAAAAACTAAAAATTGCATAAGTGATGGTAACGGCCAATTCCAAATTTAAGT -1689

-1688 TAATGCAAGTCCTACAGGCGACAAAACGTGGACGGAACATTTGAGTAGGTGCTGAAAATT -1629

-1628 GAAGGCATTGACACATTTATTCTTGGTAATTGCCAACTAACTTTGATGGAGACTAAATGG -1569

-1568 AGTCAACTCAAAACCCATATATATATCTATAAATTGTACGTCTCTGTCTCTGGGTTGAAT -1509

-1508 GTTATTGGTAAATTATTCTACTTGAATTCATCTAAATCCATTTAACATAAGCAACCAAAA -1449

-1448 TAGTGGTTATATTGTTGATTAATATTTCCTAGTAGAAGCATGCACCACCCTTACATCATG -1389

-1388 AAAATCAAGCATGCCTGAATAGGAAAATTGAAAACCATTTCTACATCCTTCAATGTATAA -1329

-1328 GAAACCAAACTTCAAATCTTTAACTAGGCCCCCCGATGATCTAAAGTCGCTTCTTTAACT -1269

-1268 TTGCCTTTGGTGTATTCTGCATCATTAGCTTTTGCTTGCACGCTCCGTATTATCATAAAC -1209

```
-1208  TCTTAATTAACTTGCATCTTTCATTTATCTCCAATCTTTACATATAAAAAGATTTCATTA  -1149
-1148  TAATATGCTAAGACTCAATAATATTACAAATGAATTGGCCAACTTTATTTAATATATAAA  -1089
-1088  AATGTCAAACAAAATAGAAATATAAATTTACGTAATATATATATATATATGTGTTTTGAA  -1029
-1028  TTATGGTAAAAGTGTTTGTATATTTCTTTTATTCTTTGATATAATTGGCTTTATTATTAT  -969
 -968  TTAAACTTGAAACTTTACCATTTTAGAAATGATTCTAGTACTAGTAAATTGAAGTTATTA  -909
 -908  ATGATTCTTTTACTTAAAAAAAGGAAAACAGAGATCTGGTAAGCCTTTCTATAAATGCGT  -849
 -848  ATTTTATGTTACTGGCTATCCTTCAATCTTGCTTTAATTATTTTCCATTATCCGTTTTTA  -789

                                                        TCA motif
 -788  TAATAGTTGTAGTAATAAGTTAATTAATTAATTAAACTTGGTCGACATAGATGTCATCTT  -729
 -728  TCAAGCATTGAAGGAAGAATATGCCAGAAATAGCGGGTGCGCGTGAAAAGCAGCAGCCCA  -669
 -668  CAGTCAAGTCTCAAGAAAGTGAACAGGTCAAGTAGCTGCAAGAGGAAGGCGGTGGACCCA  -609
 -608  TTCCAACATAACTAACCCTAGCTTACTTGTTTCCACTTTTTTTTGCTTTACATGCTTTTC  -549
 -548  GAATTCGATTTTGAGAAATAGAAAGATGTAGAAACTAAATATTTTATTTTATTTTTTAAA  -489

       ATTTCAA motif
 -488  ATTTTTTATTTCAATATTATTAATAGAACAGATAGTGTTGTGTATGTCTATTTTTTTTAT  -429
 -428  TTTAATATATAAATAACATTTTTCTTGTGTTCTAGTTATTTATATAAACAAAGTAAAAAA  -369
 -368  AGGTAATTTTTTTTTTCTTTTCATATATTTTGTTTCAAATTCTCTTCATTTATTTGTACA  -309
 -308  ACTAAGTATCCAAAGGGAATCATGAGATCATGGCAATGCCCTTCAAACAAAAGATGGTAG  -249
 -248  GGGTCCATCTGGATGTGAACTACATACATGATCCACGACTCCCTGTTTTTTTTTACTTGC  -189

                                                    CAAT box
 -188  CTTTCAATTTGTTCTCATACTCCTTTTTTTTTTCTTTTCTTTTTCCGCCAATCCTTGAAA  -129

     TATA box
 -128  CGCCTATAAATTGCAAGCCCTTTCAACTTCATTTCACACAATAGAGCTACAAAAAAAGAG  -69
  -68  AGATCGAGAACATATTGAGAGAAGTTTTATTGTTCTAGACTGAGCTAAGAAGAGATAGAG  -9
   -8  AAGAGAAAATGGAGACTTTCCCTGTTATTGACCTATCAAAGCTGAGTGGTGAAGAGAGAA  51
             M  E  T  F  P  V  I  D  L  S  K  L  S  G  E  E  R  K
   52  AACCAACCATGGAGATGATCCAAGATGCCTGTGAGAACTGGGGCTTCTTTGAGgtacgta  111
         P  T  M  E  M  I  Q  D  A  C  E  N  W  G  F  F  E
  112  ttataacaaattaattttgactgaaaacatatatatatatatatatatggaagctaatat  171
  172  attaattgtataactttgttatatagTTGGTGAACCATGGAATATCCCATGAGCTGATGG  231
                               L  V  N  H  G  I  S  H  E  L  M  D
  232  ACACTGTGGAGAGACTGACGAAGGAGCATTACAAGAAGTGTATGGAGCAAAGGTTCAAAG  291
         T  V  E  R  L  T  K  E  H  Y  K  K  C  M  E  Q  R  F  K  E
  292  AAATGGTGGCCAGTAAGGGTCTGGAGGCCGTTCAGTCCGAAATCAGTGACTTGGACTGGG  351
         M  V  A  S  K  G  L  E  A  V  Q  S  E  I  S  D  L  D  W  E
  352  AAAGCACTTTCTTTTTGCGCCACCTTCCAGTCTCCAATATGGCTGAAATCCCTGATCTTG  411
         S  T  F  F  L  R  H  L  P  V  S  N  M  A  E  I  P  D  L  D
  412  ATGAAGAATACAGgtgactacatttataatcaagtttcttttttcttctttatcttacaaa  471
         E  E  Y  R
  472  gcacaaaaggaaagatcttgtcgacttacagaagatgtttgcaaaaaaaatataatatat  531
  532  atataacagGAAGACCATGAAGGAATTTGCAGAGGAACTGGAGAAACTAGCTGAGCAACT  591
                 K  T  M  K  E  F  A  E  E  L  E  K  L  A  E  Q  L
                                        KpnI
  592  TCTGGAGGTGTTGTGCGAGAATCTTGGGTTGGAGAAAGGGTACCTGAAAAAGGCCTTCTA  651
         L  E  V  L  C  E  N  L  G  L  E  K  G  Y  L  K  K  A  F  Y
```

```
652  CGGGTCGAAAGGGCCAACCTTTGGTACAAAAGTCAGCAACTACCCTCCCTGTCCAAAACC  711
      G  S  K  G  P  T  F  G  T  K  V  S  N  Y  P  P  C  P  K  P

712  AGACCTTATCAAGGGACTGAGGGCCCACACAGATGCAGGTGGCATCATATTACTATTTCA  771
      D  L  I  K  G  L  R  A  H  T  D  A  G  G  I  I  L  L  F  Q

772  AGATGATAAGGTAAGTGGCCTCCAGCTCCTCAAGGATGGGCAATGGTTTGACGTGCCACC  831
      D  D  K  V  S  G  L  Q  L  L  K  D  G  Q  W  F  D  V  P  P

832  TATGAAACATTCCATCGTCATCAATATCGGTGACCAACTCGAGgtaataatctaacttaa  891
      M  K  H  S  I  V  I  N  I  G  D  Q  L  E

892  atttccgttaattaattaggagttgaagagaactacgatgtttttacgaacaaaacgata  951

952  actaattatgcagGTAATTACGAATGGCAAGTACAAGAGTGTGATGCACCGTGTGATTGC  1011
                  V  I  T  N  G  K  Y  K  S  V  M  H  R  V  I  A

1012 TCAAACAGATGGTACGAGAATGTCCTTAGCCTCATTCTACAACCCAGGGAGTGATGCTGT  1071
      Q  T  D  G  T  R  M  S  L  A  S  F  Y  N  P  G  S  D  A  V

1072 CATCTACCCAGCTCCCGCTTTGGTTGAGAAAGAGGCAGAGAAGTCTCAAGTGTATCCTAA  1131
      I  Y  P  A  P  A  L  V  E  K  E  A  E  K  S  Q  V  Y  P  K

1132 ATTTGTGTTCGAGGACTACATGAAGCTCTATGCAGGCCTCAAGTTCCAAGCCAAGGAGCC  1191
      F  V  F  E  D  Y  M  K  L  Y  A  G  L  K  F  Q  A  K  E  P

1192 CAGGTTTGAAGCCATGAAGGCCATGGGTCCTATTGCAACTGCTTGAGACTTATCTAGGAT  1251
      R  F  E  A  M  K  A  M  G  P  I  A  T  A  *

1252 ATTATGATCAAATCAGCTCTAGTTTGTGTTGGTGTGTGTGTGCATGTTTCTGTAATAAAA  1311

1312 ACCAATTAGGCCCTACTGATGGTGAGATTACTGTACTGATTAGGAGTGTGGAACTTTGAT  1371

                          HindIII
1372 TTGTGTAACTAAATAACTAATTATGCTGGCCCAAAGCTT  1420
```

Figure 6.4 Nucleotide sequence and deduced amino acid sequence of MeACO2.

Numbering is relative to the putative translation start site. Introns are in lower case whilst flanking regions and coding sequences are presented in upper case. Restriction sites (*Hin*dIII and *Kpn*I) used for subcloning are underlined. The putative TATA box and other *cis*-elements are in bold letters.

6.5 Discussion

6.5.1 ACO gene family in cassava

In a number of plant systems, ACC oxidase is encoded by a multigene family. Four genes were identified in tomato (Bouzayen *et al.*, 1993), four in *Petunia hybrida* (Tang *et al.*, 1993), two in mung bean (Kim and Yang, 1994) but just one in banana (Lopez-Gomez *et al.*, 1993). Whilst only a single copy was identified in banana, it is possible that banana ACC oxidase is also part of a family but clustered at a single locus (Lopez-Gomez *et al.*, 1993). The number of ACC oxidase genes in cassava was estimated by Southern analysis of cassava genomic. There were at least four bands when cassava

genomic DNA was digested by *Eco*RI, *Sac*I, *Hin*dIII and *Xba*I respectively, and was hybridized with the cDNA MeACO1 probe. This suggested that there were probably at least four ACC oxidase genes in cassava.

6.5.2 Comparison of MeACO2 and MeACO1 coding regions

The coding region sequence of MeACO2 was compared with that of cDNA MeACO1, which revealed a similarity of 79% (Figure 6.5). The deduced amino acid sequence of MeACO1 and MeACO2 was of 90% similarity and the deduced C-terminal half of the two genes were of even higher similarity (Figure 6.6).

```
MeACO2          ATGGAGACTTTCCCTGTTATTGACCTATCAAAGCTGAGTGGTGAAGAGAGAA
MeACO1          ATGGAG---TTCCCAGTCATCAACCTTGAGAAGCTTAATGGTGAGGAGAGAG
                ******   ***** ** **  ****    ***** * ****** ******

MeACO2  AACCAACCATGGAGATGATCCAAGATGCCTGTGAGAACTGGGGCTTCTTTGAGTTGGTGA
MeACO1  CTGCCACCATGGCTAAGATCAAAGATGCCTGTGAAAATTGGGGATTCTTTGAGTTGCTGA
           * *******  * **** ************* ** ***** ************ ***

MeACO2  ACCATGGAATATCCCATGAGCTGATGGACACTGTGGAGAGACTGACGAAGGAGCATTACA
MeACO1  ACCATGGAATAGAGCCAGAGTTCTTGGACAGAGTTGAGAGTATGACAAAGGGTCACTACA
        ***********   *  *** *  ******  ** *****  **** ****  ** ****

MeACO2  AGAAGTGTATGGAGCAAAGGTTCAAAGAAATGGTGGCCAGTAAGGGTCTGGAGGCCGTTC
MeACO1  GAAAATGCATGGAGCAAAGATTCAAAGAAATGGTGGCCAACAAAGGCCTCGACGCCGTCC
          ** ** ********** ******************  ** ** ** ** ***** *

MeACO2  AGTCCGAAATCAGTGACTTGGACTGGGAAAGCACTTTCTTTTTGCGCCACCTTCCAGTCT
MeACO1  AAACTGAGATCAAAGATATGGACTGGGAGAGCACCTTCTTCATCCGTCACCTCCCTGACT
        *  * ** ****  **  ********** ***** *****  * ** ***** ** * **

MeACO2  CCAATATGGCTGAAATCCCTGATCTTGATGAAGAATACAGGAAGACCATGAAGGAATTTG
MeACO1  CAAATCTTGCTCAGCTTCCTGATCTCGATGATGAACACAGGGCTGTGATGAAGGAATTTG
        * *** * *** *  * ******** ***** *** *****      **********

MeACO2  CAGAGGAACTGGAGAAACTAGCTGAGCAACTTCTGGAGGTGTTGTGCGAGAATCTTGGGT
MeACO1  CAGCAAAGCTGGAGAAACTGGCGGAGGATCTTTTGGACCTGTTGTGTGAGAATCTTGGGC
        ***   * ********** ** *** * *** ****  ******* ***********

MeACO2  TGGAGAAAGGGTACCTGAAAAAGGCCTTCTACGGGTCGAAAGGGCCAACCTTTGGTACAA
MeACO1  TCGAGAAAGGTTACCTGAAGAAGGCGTTCTATGGGTCCAGGGGTCCAACCTTTGGCACCA
        * ******** ******** ***** ***** ***** *  ** ********** ** *

MeACO2  AAGTCAGCAACTACCCTCCCTGTCCAAAACCAGACCTTATCAAGGGACTGAGGGCCCACA
MeACO1  AGGTTAGCAACTACCCACCATGTCCCAAGCCAGACCTGATCAAGGGACTCAGAGCCCACA
        * ** ********** ** ***** ** ******** ********** ** *******

MeACO2  CAGATGCAGGTGGCATCATATTACTATTTCAAGATGATAAGGTAAGTGGCCTCCAGCTCC
MeACO1  CAGACGCTGGTGGCATCATCTTGCTATTTCAAGATGACAGGGTCAGTGGCCTTCAACTCC
        **** ** ********** ** ************** * *** ******** ** ****
```

```
MeACO2  TCAAGGATGGGCAATGGTTTGACGTGCCACCTATGAAACATTCCATCGTCATCAATATCG
MeACO1  TGAAAGATGGGCAATGGATTGATGTGCCTCCTATGCGCCACTCCATTGTTGTTAACCTTG
        * ** ************ **** ***** ******   ** ***** **  * **  * *

MeACO2  GTGACCAACTCGAGGTAATTACGAATGGCAAGTACAAGAGTGTGATGCACCGTGTGATTG
MeACO1  GAGACCAGCTTGAGGTGATTACGAATGGCAAGTACAAGAGTGTCGAGCACAGAGTTGTCG
        * ***** ** ***** ***************************   **** * **  * *

MeACO2  CTCAAACAGATGGTACGAGAATGTCCTTAGCCTCATTCTACAACCCAGGGAGTGATGCTG
MeACO1  CTCAGACAGACGGTACCAGGATGTCATTAGCTTCATTCTATAACCCTGGAAGTGATGCAG
        **** ***** ***** ** ***** ***** ******** ***** ** ******** *

MeACO2  TCATCTACCCAGCTCCCGCTTTGGTTGAGAAAGAGGCAGA---GAAGTCTCAAGTGTATC
MeACO1  TGATCTACCCTGCTCCAGCTTTGGTGGAGAAAGAAGCAGAAGAGAAGAAGCAAGTCTACC
        * ******** ***** ******** ******** *****   ****   ***** ** *

MeACO2  CTAAATTTGTGTTCGAGGACTACATGAAGCTCTATGCAGGCCTCAAGTTCCAAGCCAAGG
MeACO1  CGAAATTTGTGTTTGAAGACTACATGAAGCTCTATGTTGGGTTGAAATTTCAGGCGAAGG
        * *********** ** ******************  **  * ** ** ** ** ****

MeACO2  AGCCCAGGTTTGAAGCCATGAAGGCCATGGGTCCTATTGCAACTGCTTGA  1237
MeACO1  AGCCAAGATTTGAAGCCATGAAGGCTGTGGAAAATAATGTCAATTTGGGT  950
        **** ** *****************  ***    ** **  * *    *
```

Figure 6.5 Nucleotide sequence comparison between MeACO1 and MeACO2 in coding region.

This alignment was performed using CLUSTAL W (1.8) multiple sequence alignment from http://workbench.sdsc.edu/CGI/BW.cgi.

* - single, fully conserved residue
-no consensus

```
MeACO2  METFPVIDLSKLSGEERKPTMEMIQDACENWGFFELVNHGISHELMDTVERLTKEHYKKC
MeACO1  ME-FPVINLEKLNGEERAATMAKIKDACENWGFFELLNHGIEPEFLDRVESMTKGHYRKC
        ** ****:*.**.**** .**  *:***********:****. *::* ** :** **:**

MeACO2  MEQRFKEMVASKGLEAVQSEISDLDWESTFFLRHLPVSNMAEIPDLDEEYRKTMKEFAEE
MeACO1  MEQRFKEMVANKGLDAVQTEIKDMDWESTFFIRHLPDSNLAQLPDLDDEHRAVMKEFAAK
        **********.***:***:**.*:*******:**** **:*::****:*:* .***** :


MeACO2  LEKLAEQLLEVLCENLGLEKGYLKKAFYGSKGPTFGTKVSNYPPCPKPDLIKGLRAHTDA
MeACO1  LEKLAEDLLDLLCENLGLEKGYLKKAFYGSRGPTFGTKVSNYPPCPKPDLIKGLRAHTDA
        ******:**::********************:****************************

MeACO2  GGIILLFQDDKVSGLQLLKDGQWFDVPPMKHSIVINIGDQLEVITNGKYKSVMHRVIAQT
MeACO1  GGIILLFQDDRVSGLQLLKDGQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVEHRVVAQT
        **********:***********:*****:****:*:*************** ***:***

MeACO2  DGTRMSLASFYNPGSDAVIYPAPALVEKEAE-KSQVYPKFVFEDYMKLYAGLKFQAKEPR
MeACO1  DGTRMSLASFYNPGSDAVIYPAPALVEKEAEEKKQVYPKFVFEDYMKLYVGLKFQAKEPR
        ******************************* *.***************.**********
```

```
MeACO2   FEAMKAMGPIATA------------
MeACO1   FEAMKAVENNVNLGPNCYCLIINYY
         ******: ..
```

Figure 6.6 comparison of deduced amino acid sequence of MeACO1 and MeACO2
This alignment was performed using CLUSTAL W (1.8) multiple sequence alignment from http://workbench.sdsc.edu/CGI/BW.cgi.

* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
- no consensus

6.5.3 Comparisons of ACO genes isolated from cassava and other plants

6.5.3.1 Comparison of the deduced amino acid sequences

ACC oxidase genes from different plant species are of high similarity both in nucleic acid sequence and deduced amino acid sequence. Cassava MeACO2 had 85% similarity to the ACC oxidase gene from *Prunus persica* in the coding region (Tang *et al.*, 1993). The isolated cassava ACO genes including cDNA MeACO1 and MeACO2 shared very high amino acid sequence identity with a number of ACO genes from other plants such as *Carica papaya* (84% identity and 88% similarity, GeneBank accession number: U68215), *Actinidia deliciosa* (83% and 89%, M97961), *Prunus persica* (81% and 89%, AF129074), petunia (81% and 86%, Q08507), tomato (81% and 89%, X58273), *Populus euramericana* (81% and 87%, AB033504) and *Nicotiana glutinosa* (79% and 88%, U54565). To demonstrate the high similarity, the deduced amino acid sequences of the two cassava genes were compared with four tomato ACO genes, a well-characterised ACO gene family (Fig 6.7). As can be seen, the deduced amino acid sequences of these ACO genes were highly conserved from the N terminal to the C terminal except for the two short regions near the C terminal. The identities between MeACO2 and tomato ACO genes including LeACO1, LeACO2, LeACO3 and LeACO3 were 83%, 79%, 82% and 81% respectively. This means that MeACO2 was much more similar to tomato LeACO1 and LeACO3 than to MeACO1 (81% identity with MeACO2). Tomato LeACO1 and LeACO3 are similar in expression patterns. Both were expressed in fruits at the start of ripening and leaves undergoing senescence, although the accumulation level of the LeACO1 mRNA was higher than that of the LeACO3 mRNA (Barry *et al.*, 1996). LeACO1 was also strongly induced by

mechanical wounding (Barry *et al.*, 1996). The high similarity between MeACO2 and these tomato ACO genes may suggest MeACO2 is involved in wound response such as PPD and senescence.

```
MeACO2          MET-FPVIDLSKLSGEERKPTMEMIQDACENWGFFELVNHGISHELMDTVERLTKEHYKK
MeACO1          ME--FPVINLEKLNGEERAATMAKIKDACENWGFFELLNHGIEPEFLDRVESMTKGHYRK
LeACO1          MEN-FPIINLEKLNGDERANTMEMIKDACENWGFFELVNHGIPHEVMDTVEKMTKGHYKK
LeACO2          MEN-FPIINLEKLNGAERVATMEKINDACENWGFFELVNHGIPHEVMDTVEKLTKGHYKK
LeACO3          MEN-FPIINLENLNGDERAKTMEMIKDACENWGFFELVNHGIPHEVMDTVEKLTKGHYKK
LeACO4          MESNFPVVDMGLLQTEKRPEAMDKIKDACENWGFFELVNHGISHELLDAVENLTKGHYKK
                **  **:::: *.  :*  :*  *:***********:****  *.:* ** :** **:*

MeACO2          CMEQRFKEMVASKGLEAVQSEISDLDWESTFFLRHLPVSNMAEIPDLDEEYRKTMKEFAE
MeACO1          CMEQRFKEMVANKGLDAVQTEIKDMDWESTFFIRHLPDSNLAQLPDLDDEHRAVMKEFAA
LeACO1          CMEQRFKELVASKGLEAVQAEVTDLDWESTFFLRHLPTSNISQVPDLDEEYREVMRDFAK
LeACO2          CMEQRFKELVAKKGLEGVEVEVTDMDWESTFFLRHLPSSNISQLPDLDDVYREVMRDFRK
LeACO3          CMEQRFKELVASKGLEAVQAEVTDLDWESTFFLRHLPTSNISQVPDLDEEYREVMRDFAK
LeACO4          CMEQRFKEMVASKGLEAVQTEIDDLDWESTFFLKHLPVSNVYEVPDLDDEYRKVMKDFAL
                ********:**.***:.*: *: *:*******::*** **: ::****: :* .*::*

MeACO2          ELEKLAEQLLEVLCENLGLEKGYLKKAFYGSKGPTFGTKVSNYPPCPKPDLIKGLRAHTD
MeACO1          KLEKLAEDLLDLLCENLGLEKGYLKKAFYGSRGPTFGTKVSNYPPCPKPDLIKGLRAHTD
LeACO1          RLEKLAEELLDLLCENLGLEKGYLKNAFYGSKGPNFGTKVSNYPPCPKPDLIKGLRAHTD
LeACO2          RLEKLAEELLDLLCENLGLEKSYLKNTFYGSKGPNFGTKVSNYPPCPKPDLIKGLRAHTD
LeACO3          RLEKLAEELLDLLCENLGLEKGYLKNAFYGSKGPNFGTKVSNYPPCPKPDLIKGLRAHTD
LeACO4          KLEKLAENLLDLLCENLGLEKGYLKKAFYGSKGPTFGTKVSNYPPCPKPDLIKGLRAHTD
                .******:**::*********.***::****:**.************************

MeACO2          AGGIILLFQDDKVSGLQLLKDGQWFDVPPMKHSIVINIGDQLEVITNGKYKSVMHRVIAQ
MeACO1          AGGIILLFQDDRVSGLQLLKDGQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVEHRVVAQ
LeACO1          AGGIILLFQDDKVSGLQLLKDEQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVLHRVIAQ
LeACO2          AGGIILLFQDDKVSGLQLLKDGRWIDVPPMRHSIVVNLGDQLEVITNGKYKSVMHRVIAQ
LeACO3          AGGIILLFQDDKVSGLQLLKDEQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVMHRVIAQ
LeACO4          AGGIILLFQDDKVSGLQLLKDGNWIDVPPMKHSIVINLGDQLEVITNGRYKSIEHRVIAQ
                ***********:********* .*:*****:****:*:**********:***: ***:**

MeACO2          TDGTRMSLASFYNPGSDAVIYPAPALVEK--EAEKSQVYPKFVFEDYMKLYAGLKFQAKE
MeACO1          TDGTRMSLASFYNPGSDAVIYPAPALVEKE-AEEKKQVYPKFVFEDYMKLYVGLKFQAKE
LeACO1          TDGTRMSLASFYNPGSDAVIYPAKTLVEKEAEES-TQVYPKFVFDDYMKLYAGLKFQAKE
LeACO2          KDGTRMSLASFYNPGNDALIYPAPALVDKEAEEHNKQVYPKFMFDDYMKLYANLKFQAKE
LeACO3          TDGTRMSLASFYNPGNDAVIYPAPSLIE----ES-KQVYPKFVFDDYMKLYAGLKFQPKE
LeACO4          QDGTRMSIASFYNPGSDAVIFPAPELIEKT-EEDIKLKYPKFVFEDYMKLYAGLKFQAKE
                 ******:*******.**:*:**  *::       .  ****:*:******..****.**

MeACO2          PRFEAMKA-------MGP-----IATA-
MeACO1          PRFEAMKAVENNVN-LGPNCYCLIINYY
LeACO1          PRFEAMKAMES-----DP-----IASA-
LeACO2          PRFEAMKAMES-----DP-----IAIA-
LeACO3          PRFEAMKAMEANVELVDQ-----IASA-
LeACO4          PRFEAMKAVETTVN-LGP-----IETV-
                ********         .      *
```

Figure 6.7 Comparison of the deduced amino acid sequences of MeACO1 and MeACO2 with those of four tomato ACC oxidase genes.

This alignment was performed using CLUSTAL W (1.8) multiple sequence alignment from http://workbench.sdsc.edu/CGI/BW.cgi. Accession number for tomato ACO genes in GeneBank, LeACO1: X58273, LeACO2: Y00478, LeACO3: Z54199 and LeACO4: AB013101

* - single, fully conserved residue

: - conservation of strong groups

. - conservation of weak groups

- no consensus

6.5.3.2 Promoter and *cis*-element analysis

Not only was high homology observed in the coding region among ACO genes, similar motifs or putative regulatory elements were also identified in the 5' flanking regions of ACC oxidase genes in different plant species. The TCA motif and near matches have been shown to be present in over 30 stress-induced genes including glucanases, PR genes and wound-induced genes (Goldsbrough *et al.,* 1993). The TCA motif, TCATCTTCTT, or near matches, have been identified in 5' flanking region of tomato ACO3 (Barry, 1996) and MeACO2 (at –723 to –733 see Figure 6.4), indicating their possible regulation by stresses. The motif ATTTCAAA or AATTCAAA was identified to be related to ethylene inducibility from two ethylene-regulated genes, E4 gene in tomato and glutathione-S-transferase (GST1) gene in carnation (Montgomery *et al.,* 1993; Itzaki *et al.,* 1994). An ethylene inducible motif ATTTCAA was also identified in the 5' flanking region of tomato ACO3 (Barry, 1996), mung bean ACO gene (Jin *et al.,* 1999) and the cassava MeACO2 (-471 to –465, -2127 to –2121 and –2431 to –2424 see Figure 6.4). It was shown by Barry (1996) that ACO3 was inducible in green fruit by ethylene. ACO3 transcripts were detectable only 2 h after ethylene treatment and continued to increase to a high level during the 8 hours experiment of ethylene treatment. The mung bean ACO gene containing the ethylene inducible motif was also shown to be inducible by ethylene. Therefore, it is likely that MeACO2 is ethylene inducible.

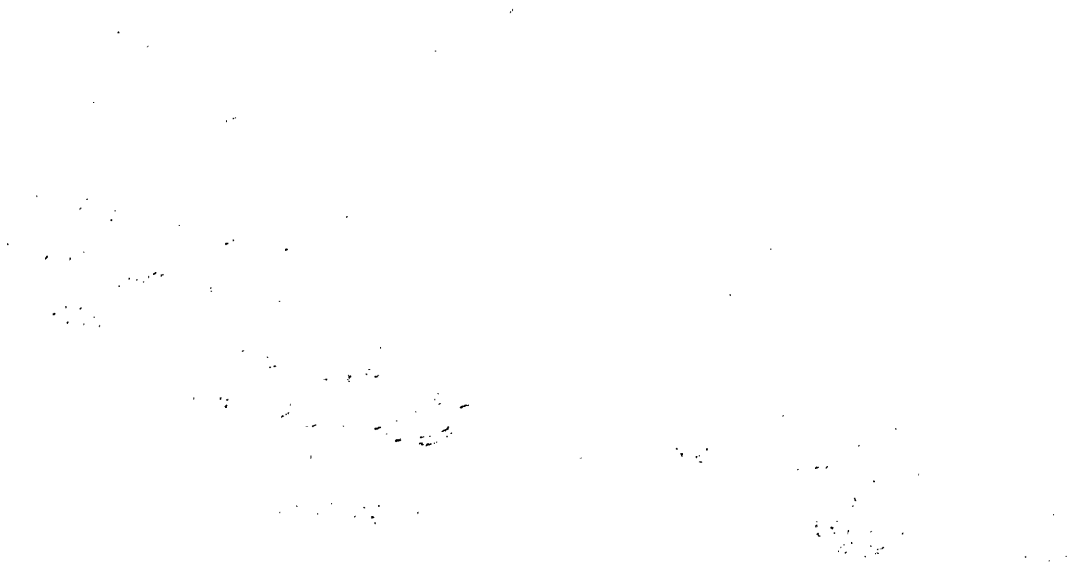
The sequence flanking the putative TATA box of MeACO2 was similar to those of tomato ACO genes (see Figure 6.8). It was also interesting to note that 9 bp downstream of the putative TATA box in MeACO2 are identical to 9 bp sequence downstream of TATA box in LeACO1 and 3 (The genomic sequence of LeACO4 was not available in Genbank). This enabled the MeACO2 TATA box to be identified with confidence.

```
MeACO2 AACGCCTATAAATTGCAAGCCCTTTCAACTTCATTTCACAC
LEACO3 TTCTCCTATAAATA      CCCTTTCAAAGAATCACTCTTT
LEACO1 TTCTACTATAAATA      CCCTTTCAAAGCCTCATTATTT
LEACO2 TTTCCCTATAAATA      CCCAAACAAAGCCTCAATCTTT
           **********      **********
```

Figure. 6.8 Optimal sequence alignment of the TATA region of MeACO2 and tomato ACO genes. The identical nucleotides are marked with *. Putative TATA box is in bold.

6.5.3.3 Comparison of introns in ACO genes

In addition to the high similarity in nucleotide sequences and amino acid sequence, ACC oxidase genes from different species also shared similar structures. For example, ACC oxidase genes in banana, *Brassica napus*, melon, petunia and tomato had four exons interrupted by three introns and shared the same intron-exon junction sequences and similar location of introns (Lopez-Gomez *et al.*, 1993; Bouzayen *et al.*, 1993). Similarly, cassava MeACO2 also consisted of four exons and three introns. Table 6.1 shows the sizes and junctions of the three introns from MeACO2 and tomato ACO genes. Though the size of each intron from these genes varied, the sequences at the junction of exon-intron or intron-exon were highly conserved. For example, the sequences of the 3' end of the first exon were identical between MeACO2 and three tomato ACO genes, which were TTTGAG. Similarly sequence TTGGTG from the 5' end of the second exon was shared among these genes. MeACO2 and LeACO1 appeared to be similar in the sizes of the 2nd and 3rd intron.



Gene	Intron	Length (bp)	Donor Acceptor
MeACO2	1	94	TTTGAGgtacgt...atatatTTGGTG
LeACO1		278	TTTGAGgtaatc...ttggagTTGGTG
LeACO2		334	TTTGAGgtacaa...gtatagTTGGTG
LeACO3		225	TTTGAGgtaata...ttgtagTTGGTG
MeACO2	2	117	ATACAGgtgact...taacagGAAGAC
LeACO1		106	ATACAGgtacat...atgcagAGAGGT
LeACO2		825	ATATAGgtaaac...atgtagGGAAGT
LeACO3		106	ATACAGgtacat...atgcagAGAGGT
MeACO2	3	90	CTCGAGgtaata...atgcagGTAATT
LeACO1		89	CTTGAGgtacaa...aaacagGTGATC
LeACO2		398	CTCGAGgtaaaa...aaacagGTAATC
LeACO3		346	CTTGAGgtacaa...aaacagGTGATT

Table 6.1 Intron splice junction sites and intron lengths of the MeACO2 and comparison between MeACO2 and three tomato ACC oxidase genes.

Intron sequences are represented in lower case whilst exonic sequences are in upper case. The splice junction sites conform to the consensus plant sequences (Brown, 1986). The genomic sequence of LeACO4 was not available in Genebank.

Comparison between cassava MeACO2 and tomato ACO genes seems to suggest that MeACO2 is a cassava homologue of LeACO1. MeACO2 shared higher identity of amino acid sequence with LeACO1 than with other tomato genes. These two genes also had the same intron-exon structure and they were almost identical in the size of two introns. MeACO2 may be ethylene inducible as suggested by the presence of a ethylene inducible motif in its promoter region and LeACO1 has been shown to be induced by ethylene (Barry, 1996). In tomato, the transcripts of three ACO genes were accumulated to varying levels in different tissues at various stages of development. LeACO1 was at a low level in a variety of organs but increased dramatically at the start of fruit ripening and senescence of leaves and flowers, and also in response to mechanical wounding of leaves. LeACO2 was only detectable in flowers and during seed germination. LeACO3 was similar to LeACO1 but with substantially lower expression level and was not induced by wounding (Barry, 1996). During ripening,

LeACO4 expression increased in a similar trend to LeACO1 but the expression level was lower (Nakatsuka *et al.,* 1998). Expression patterns of MeACO2 are needed to clarify whether it is the cassava homologue of LeACO1, which can be induced by mechanical wounding and ethylene.

6.6 Summary

Southern hybridization showed that ACC oxidase in cassava was encoded by a multigene family. An ACO genomic clone, MeACO2, was isolated from a cassava genomic library using cMeACO1 as a probe. This clone was not the corresponding genomic clone of cMeACO1, but it had high similarity to MeACO1 cDNA at both nucleotide and deduced amino acid sequences. The MeACO2 was 6349 bp in length, consisting of 4928 bp 5' UTR or 5' flanking region and four putative exons interrupted by three introns. The deduced amino acid sequence was 312 amino acids. The clone was compared with a well-characterised ACO gene family, four tomato ACO genes, which showed that it was similar to this group of genes in sequence and structure, especially to LeACO1. In the promoter region, MeACO2 shared a similar sequence around TATA box with three tomato ACO genes. MeACO2 and three tomato genes all had four exons separated by three introns, and the locations of the three introns in all these genes were similar. It was also noted that these genes shared highly similar sequences in the junctions of exon/intron. It was speculated that MeACO2 might be a cassava homologue of tomato LeACO1 and might be regulated in a similar way such as by mechanical wounding and ethylene.

CHAPTER SEVEN: GENERAL DISCUSSION AND FUTURE WORK

This thesis described the isolation of stress-related genes MePAL1, MePAL2 and MeACO2 from a cassava genomic library, the comparison of those genes with homologous genes from other plants and the investigation of the properties of MePAL2 during cassava development and PPD. The expression patterns of MePAL2 promoter-GUS fusions in both transgenic cassava and rice were analyzed. To the best of my knowledge, this is the first successful example of the functional test of a cassava promoter in transgenic cassava, though the expression of heterologous promoter fusions has been analyzed in transgenic cassava. The success demonstrates that the cassava transformation system using particle bombardment has reached mature stage and can be used to analyse gene function and for genetic modification of cassava. This study is also the first example of dicot PAL promoter being expressed in a monocot plant, rice. The data support that the mechanisms for regulation of PAL gene expression during plant development are conserved between dicots and monocots.

In this chapter the possible roles of the known PAL genes in PPD, the properties and possible application of the MePAL2 promoter, and future work related to these genes and ACC oxidase are discussed. The future of cassava research is also discussed.

7.1 Cassava PAL genes and their possible roles during PPD

PAL is encoded by a multigene family in cassava. Up to now, four PAL genes have been isolated from cassava, though for two of them (MePAL and MePAL3), only partial sequences are known. The similarities of those four cassava PAL gene were quite high, around 79%-98%, especially MePAL2 and MePAL3 with only 15 nucleotide differences over 1000 bp (for detailed discussion, see Chapter 3). MePAL1 was isolated from PPD related cassava cDNA library (Han, 2000). MePAL1 mRNA was not detected in freshly harvested tubers but detected by RT-PCR in the cassava

tuber eight hours after harvest. This may indicate that MePAL1 was directly related to PPD-specific induction in tuber (Han, 2000). MePAL was a PCR fragment amplified from cassava genomic DNA using degenerate primers and showed PPD-induced expression pattern, which was similar to that of MePAL1 (Periera *et al.,* 1999). These two cassava PAL genes, MePAL1 and MePAL, may be closely associated with vascular discoloration. MePAL3 were isolated from a PPD related cassava cDNA library (Han, 2000). Though there is no expression data for MePAL3, it might be the same expression pattern as MePAL2 since its sequences were 98% similar to MePAL2. In parsley, PAL1 and PAL2, which had almost identical sequences throughout the regions analyzed, showed the same expression pattern in the cases tested, such as under light, elicitor and wounding stress (Logemann *et al.,* 1995). Though there was no direct evidence to suggest that MePAL2 was involved in PPD, the GUS expression pattern driven by the 840 bp of MePAL2 promoter showed that the promoter was active in xylem parenchyma of the cassava tuber, in where the initial symptoms of PPD occur. The MePAL840 promoter was not only active in vascular tissues of freshly-harvested tuber but also enhanced during the development of PPD. Furthermore, pruning, which can delay the development of PPD (Tanaka *et al.*, 1984), postponed the appearance of peak activity of the MePAL2 promoter and also reduced the peak activity as indicated by GUS activity (Figure 5.12). These observations suggest that MePAL2 expression is related to PPD. It is especially interesting to note that in the tuber harvested from three months old plants the promoter activity peaked 24 hr post-harvest, which is the time point that the content of scopoletin peaked as observed by Wheatley and Schwabe (1985). However, it was not clear whether the enhanced activity of the MePAL2 promoter during PPD would contribute to PAL activity related to the production of phenolic compounds linked to vascular discoloration. Since scopoletin probably plays a role in PPD, as it was demonstrated that exogenous application of scopoletin promoted PPD significantly and delayed PPD by pruning was related to low scopoletin content (Wheatley and Schwabe, 1985), it is important to investigate the changes in scopoletin content in the transgenic cassava tuber during PPD and compare the trend with that of the PAL promoter activity, which may indicate whether MePAL2 is linked with the production of scopoletin during PPD.

Even though PAL genes have been studied in many plant species, only a few PAL gene families have been analysed, and none of the known PAL gene families has been studied by analysing the promoter of each family member. Since phenolic compounds seem to play an important roles in PPD, one of the critical constraints for cassava production, it is necessary to investigate the PAL gene family in detail in order to gain an insight into which PAL gene is more important in PPD. Therefore, analysis of the other PAL promoters using the same approach as for MePAL2 should be carried out and the expression patterns of these promoters during PPD compared with each other, to the pattern of discoloration and accumulation of scopoletin. Alternatively, these PAL genes may be over-expressed by introducing strong promoter-driven sense PAL genes or down-regulated by sense-suppression or antisense techniques to observe their effect on PPD in order to establish how the PAL genes are linked to the PPD.

7.2 The potential properties of MePAL2 promoter

In this research, the 5' flanking region of MePAL2 was isolated, in which a number of putative regulatory elements and motifs were identified. The 5' flanking region consists largely A/T-rich and G/C- or C rich regions. In addition to the characteristics of eukaryotic promoters such as TATA box and CCAAT box, the MePAL2 promoter also has motifs with identical or similar sequences to stress inducible elements. These include wound inducible element G box and H box, jasmonic acid responsive elements, P-A-L boxes (or AC boxes) inducible by elicitor and light in vivo, auxin and ethylene inducible elements. P-A-L boxes in the bean PAL promoter were shown to be involved in conferring xylem tissue-specific expression (Hotton *et al.,* 1995). The GUS expression pattern driven by the 840 bp (containing the P-A-L box) of MePAL2 promoter in transgenic cassava showed that the promoter was active in the xylem parenchyma, as the other PAL promoters with P-A-L boxes analysed in other plants (Gray-Mitsumune *et al.*, 1999). It has been shown in this research that the activities of the promoter were enhanced during PPD, which may be related to the putative wound inducible elements in the promoter. Further experiments should be carried out to test whether the promoter can be enhanced in other tissues such as leaves and stems upon wounding. The MePAL2 promoter may also be inducible by elicitors, as it was shown that the P-A-L boxes were necessary for elicitor- or light- mediated PAL gene

activation. In transgenic tobacco containing the rice ZB8 PAL promoter-GUS, the GUS activity increased five to ten fold in tobacco leaves infected with tobacco mosaic virus (TMV), compared to the GUS activity in uninfected tobacco leaves (Zhu *et al.,* 1995). The response of the MePAL2 promoter to pathogen attack is being investigated by analysing the GUS activity of infected leaves and the expression pattern at the infected sites using the PAL840GUS (Kemp B, Personal communication).

In order to further understand the properties of the promoter, promoter deletions PAL400, PAL260 and PAL200 based on distribution of putative elements were fused to GUS reporter gene and transferred into embryogenic cassava suspension cells to obtain transgenic plants to test the stable activities of these truncated promoters. At the callus stage, PAL400 showed similar GUS expression to PAL840. In transgenic rice plants, the expression patterns of GUS driven by PAL840 and PAL400 were the same, indicating that about half of the promoter sequence by the 5' end of PAL840 was not vital for tissue specific expression, but essential for the full strength of the promoter as indicated by the higher transient GUS expression directed by PAL840 than PAL400. Unfortunately, no transgenic cassava plants were obtained from PAL400, PAL260 and PAL200 by the time this PhD project was terminated. Some lines are being produced from PAL400-GUS but are still at the callus stage. Obtaining transgenic cassava lines from these constructs and analysis of these lines would help to confirm the role of the putative P-A-L boxes in tissue specific expression and in response to elicitor and UV light. As the putative JA responsive element is truncated in PAL400, it would be interesting to investigate and compare the responses of the PAL840 and PAL400 promoter in leaves to exogenous applied JA to test the fidelity of the element.

7.3 The possible application of MePAL2 promoter in genetic engineering of cassava

With increasing population pressures that require greater productivity, cassava may become one of the most important crops of the 21st century (Taylor *et al.,* 1999). However, traditional breeding systems are hindered due to the highly heterozygous nature and asynchronous flowering of cassava. The emergence of biotechnological tools provide opportunities to overcome these problems.

Since the significant breakthroughs in cassava transformation in 1996 (Schopke *et al.*, 1996; Li *et al.*, 1996), attempts have been made to genetically improve the crop for resistance to virus or herbicide, modified starch and reduction in toxicity (Sarria *et al.*, 2000, Arias-Garzon *et al.*, 1998, Munyikwa *et al.*, 1998, Taylor *et al.*, 1998). So far all the transgenic cassava plants obtained used constitutive promoters such as CaMV 35S or cassava vein mosaic virus (CsVMV) promoters to drive the transgenes. In order to express transgenes more efficiently and specifically, tissue specific and inducible promoters are required.

In the present research, transformation was used to analyse the PAL promoter in cassava, which is the first example of studying an endogenous promoter using transgenic cassava. The GUS expression patterns driven by 840 bp of MePAL2 promoter suggests tissue-specific properties of the promoter; the promoter-GUS reporter gene was expressed in the xylem parenchyma of leaves, petioles, stems, and roots. Expression data from a constitutive promoter 35S driven GUS would be informative in understanding how specific the MePAL2 promoter is compared to the commonly used promoter. Therefore, the MePAL2 promoter has the potential to be used in genetic engineering to drive pathogen resistance genes in the xylem to enable the plant to become resistant to xylem pathogens or pathogens spreading through xylem. *Xa21*, isolated from a wild relative of rice and known for its resistance to bacterial leaf blight caused by *Xanthomonas oryzae* pv. *oryxae*, was transferred to rice and fifty lines of transgenic rice plants carrying the cloned *Xa21* gene displayed high levels of resistance to the pathogen (Song *et al.*, 1995; Ronald 1997). Cassava production is also affected by bacterial blight (CBB) caused by *Xanthomonas axonopodis* pv. *manihotis*. The pathogen penetrates the host through stomata and then spreads systemically to the vascular tissues in all parts of the plant, most of this movement occurs through the xylem vessels. Five cassava BAC clones containing homologues to the *Xa21* were identified by screening a cassava BAC library with a probe made from *Xa21* at CIAT (Cited as personal communication with M Fregene by Taylor, 2000). The MePAL2 promoter may be used to direct specific expression of either rice *Xa21* or its cassava homologue to the xylem of cassava, via which the CBB pathogen was spread, in order to genetically manipulate CBB.

7.4 To dissect the roles of ACC oxidase genes in cassava using gene silencing and their application to control PPD

Two ACC oxidase genes, MeACO1 (by Han, 2000) and MeACO2 (present research), have been isolated from cassava. However, no expression data were obtained except that MeACO1 was expressed during PPD. The roles of these genes in cassava development, and particularly in the development of PPD, could be analysed using sense or antisense suppression techniques. Suppression of target endogenous genes by introducing an antisense transgene has been widely used in aspects such as identifying gene function (i.e. tomato pTOM13) (Hamilton *et al.*, 1990), crop improvement (i.e. prolonged shelf-life of tomato) (Hamilton *et al.*, 1990) and studies of metabolic pathways (i.e.chalcone synthase in *Petunia*) (Bourque, 1995). Similarly, more and more evidence demonstrate that the introduction of sense transgenes into the plant genome may also lead to silencing of endogenous genes or resistance to viruses if the transgene is homologous to a virus gene (Napoli *et al.*, 1990; Smith *et al.*, 1990; Baulcombe, 1999). The advantage of transferring a sense transgene into plants is that two types of transgenic plants may be obtained, one with the target gene being down-regulated and the other with over-expressing genes. These two sets of plants would help to identify the effect of endogenous gene or /and transgene on plant development. For example, sense MeACO1 or MeACO2 driven by CaMV 35 S promoter could be introduced into cassava to study the effect of ethylene on PPD. Suppression of endogenous ACO gene(s) may reduce the ethylene biosynthesis or may not have much effect on ethylene production in the deteriorating tuber if the expression of these genes were low and they were not the main contributor to ethylene biosynthesis in tuber. Over-expressing the ACO gene could lead to an increase in ethylene biosynthesis, which could promote the development of PPD, if ethylene plays a role in the process.

PPD is a special wounding response without much healing process or an un-localized wounding response (Beeching *et al.*, 1995). It could be that the harvested tuber is too sensitive to mechanical wounding and that the wounding signals are transmitted through the vascular tissues rather than being localised to the wound site, causing responses (represented by the vascular discoloration) throughout the vascular system. It

was shown that more ethylene was produced during PPD in a susceptible cultivar than in a cultivar with less susceptibility (Hirose, 1986). If this hypothesis is true, it could be possible that down-regulation of ACO genes by expressing a sense or antisense ACO transgene may result in transgenic lines less susceptible to PPD. In tomato, antisense or sense suppression of ACO1 led to delayed ripening due to the reduction of ethylene production (Hamilton *et al.*, 1990 and 1998).

7.5 Fast forward cassava biology by a molecular approach

Molecular techniques have been extremely useful tools for understanding plant biology. There is no doubt that application of these techniques in cassava research would not only promote our knowledge of the crop but also fulfil the potential to improve crop production using these techniques. In this research, the transformation technique was successfully used for analysing the PAL promoter in cassava. Even though the transformation system is still time- and labour-consuming compared to the transformation systems of model plants such as *Arabidopsis*, tobacco and tomato, the successful application of the technique in this research not only demonstrated the maturity of the technique but also indicated the promising impact of the technique on the progress in understanding cassava at molecular level. With the dramatic increase in the number of plant genes being isolated, especially the complete sequence of the *Arabidopsis* genome, transgenic plants are playing increasingly important roles in characterising functions of genes and understanding plant development. In Genebank of NCBI, 76 cassava genes are registered, which is a very small number compared to the huge Genebank database and its rate of expansion. However, more and more genes are being isolated from cassava. A number of cDNA libraries, genomic libraries and a cassava BAC library have been constructed and genes are being identified from these libraries. Using cDNA-AFLP (amplified fragment length polymorphism), more than 500 transcript-derived fragments from cassava were obtained and 50 of them were cloned and sequenced, most of which showed no homology to known genes in Genebank (Suarez *et al.*, 2000). The cassava transformation system would definitely accelerate the pace of characterising these genes and their functions.

It is also encouraging that the cassava biotechnologies have been being transferred to developing countries where cassava is the staple crop. For example, International Laboratory of Tropical Agricultural Biology (ILTAB, Danforth Plant Centre, USA) has been training scientists from tropical countries in cassava transformation technique. The technology transfer would help these countries to develop their own expertises to address their specific constraints in cassava production using biotechnology.

REFERENCES

Abeles FB, Morgan PW, Saltveit ME (1992) Ethylene in plant biology. 2nd edition. Academic Press, New York. 414 pp

Aida R, Yoshida T, Ichimura K, Goto R and Shibata M (1998) Extension of flower longevity in transgenic torenia plants incorporating ACC oxidase transgene. Plant Science 138(1): 91-101

Appert C, Logemann E, Hahlbrock K, Schmid J, Amrhein N (1994) Structural and catalytic properties of the 4 phenylalanine ammonia-lyase isoenzymes from parsley (petroselinum-crispum nym). European Journal of Biochemistry 225 (1): 491-499

Arias-Garzon DI and Sayre RT (1993) Tissue specific inhibition of transient gene expression in cassava (*Manihot esculenta* Crantz). Plant Science. 93(1-2): 121-130

Arias-Garzon DI and Sayre RT (1998) Genetic engineering approaches to reducing the cyanide toxicity in cassava (*Manihot esculenta*, Crantz). The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Averre CW (1967) Vascular streaking of stored cassava roots. Proceedings of 1st International Symposium on Tropical Root Crops. Trinidad 2(4): 31-35

Awoleye F, van Duren M, Dolezel J and Novak FJ (1994) Nuclear DNA content and *in vitro* induced somatic polyploidization cassava (*Manihot esculenta* Crantz) breeding. Euphytica 76: 195-202

Awoyinka AF, Abegunde VO and Adewusi SRA (1995) Nutrient content of young cassava leaves and assessment of their acceptance as a green vegetable in Nigeria. Plant Food for Human Nutrition 47: 21-28

Bach M and Seitz HU (1997) Elicitor-induced defence responses of a suspension-cultured woody Plant (*Larix decidua*) and possible mechanisms of signal transduction. Canadian Journal of Botany 75: 1243-1251

Balague C, Watson CF, Turner AJ, Rouge P, Picton S, Pech JC, Grierson D (1993) Isolation of a ripening and wound-induced cDNA from *Cucumis melo* L. encoding a protein with homology to the ethylene-forming enzyme. European Journal of Biochemistry 212(1):27-34

Barry CS, Blume B, Bouzayen M, Cooper W, Hamilton AJ, Grierson D (1996) Differential expression of the 1-aminocyclopropane-1-carboxylate oxidase gene family of tomato. Plant Journal 9(4): 525-535

Baulcombe DC (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. Plant Molecular Biology, 32: 79-88.

Baulcombe DC (1999) Fast forward genetics based on virus-induced gene silencing. Current Opinion in Plant Biology 2: 109-113.

Baulcombe DC (1999) Viruses and gene silencing in plants. Archives of Virology Supplement 15: 189-201

Beeching JR, Dodge AD, Moore KG and Wenham JE (1995) Physiological deterioration in cassava: An incomplete wound response? In: Thro AM & W Roca (Eds.), The Cassava Biotechnology Network: Proceedings of the Second International Scientific Meeting. CIAT Cali, Colombia. 729-736 pp

Beilmann A, Pfitzner AJP, Goodman HM and Pfitzner UM (1991) Functional analysis of the pathogenesis-related 1a protein gene minimal promoter region. Comparison of reporter gene expression in transient and in stable transfections. European Journal of Biochemistry 196: 415-421

Bell-Lelong DA, Cusumano Jc, Meyer K, Chapple C (1997) Cinnamate-4-hydroxylase expression in Arabidopsis:regulation in response to development and the environment. Plant Physiology 113:729-738

Bennet RN and Wallsgrove RM (1994) Secondary metabolites in plant defence mechanisms. New Phytology 127: 617-633

Bergey DR; Orozco-Cardenas M; de Moura DS; Ryan CA (1999) A wound- and systemin-inducible polygalacturonase in tomato leaves. Proceedings of the National Academy of Sciences of USA. 96: 1756-1760

Bernards MA and Lewis NG (1998) The macromolecular aromatic domain in suberized tissue: a changing paradigm. Phytochemistry 47(6): 915-933

Bevan M, Shufflebottom D, Edwards K, Jefferson R, Schuch W (1989) Tissue- and cell-specific activity of a phenylalanine ammonia-lyase promoter in transgenic plants EMBO Journal 8:1899-1906

Bevan MW, Flavell RB and Chilton MD (1983) A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. Nature 304: 184-187.

Bleecker AAB, Estelle MA, Somerville C and Kende H (1988) Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. Science 241: 1086-1089

Blume B, Barry CS, Hamilton AJ, Bouzayen M, Grierson D (1997) Identification of transposon-like elements in non-coding regions of tomato ACC oxidase genes. Molecular and General Genetics 254 (3): 297-303

Blume B, Grierson D(1997) Expression of ACC oxidase promoter-GUS fusions in tomato and *Nicotiana plumbaginifolia* regulated by developmental and environmental stimuli. Plant Journal 12(4):731-46

Boher B and Verdier (1994) Cassava bacterial blight in Africa: the state of knowledge and implications for designing control strategies. African Journal of Root and Tuber Crops 2:505-509

Bolhuis GG (1966) Influence of length of the illumination period on root formation in cassava. Netherlands Journal of Agricultural Science.14, 151-262

Bohl-Zenger S, Potrykus I, Puonti-Kaerlas J (1998) Isolation of root specific promoters from cassava. Brazilian Cassava Journal, Supplement, Fourth International Meeting Cassava Biotechnology Network, Salvador, Brazil, pp 201-205

Bokanga M (1994) Distribution of cyanogenic potential in cassava germplasm. Acta Horticulturae 375:117-123

Bokanga M; Ekayanake IJ; Dixon AGO and Porto CM (1994) Genotype environment interactions for cyanogenic potential in cassava. Acta Horticulturae 375: 131-140

Boller T (1990) Ethylene and plant-pathogen interactions. In: Current Topics in Plant Physiology: An American Society of Plant Physiologists Series, Vol. 5. Polyamines and Ethylene: Biochemistry, Physiology, and Interactions. Ed by Flores HE, Arteca RN and JC Shannon. American Society of Plant Physiologists: Rockville, Maryland, USA. Illus. Paper. ISBN 0-943088-19-4. pp138-145.

Booth RH (1976) Storage of fresh cassava I. Post-harvest deterioration and its control. Experimental Agriculture 12:103-111

Booth RH (1977) Storage of fresh cassava (*Manihot esculenta*) II. Simple storage techniques. Experimental Agriculture 13: 119-128

Boudet AM (1998) A new view of lignification. Trends in Plant Science 3(2): 67-71

Boudet AM, Lapierre C and Grima-Pettenatti J (1995) Biochemistry and molecular biology of lignification. New Phytology 129: 203-236

Bourque JE (1995) Antisense strategies for genetic manipulations in plants. Plant Science 105: 125-149

Bouzayen M, Cooper W, Barry CS, Zegzouti H, Hamilton AJ Grierson D (1993) EFE multigene family in tomato plant: expression and characterisation. In:Cellular and Molecular Aspects of the Plant Hormone Ethylene. Eds. Pech JC, Latche A, Balague C.pp 76-81. Kluwer Academic Publishers.

Bowles DJ (1990) Defense-related proteins in higher plants. Annual Review Biochemistry 59: 873-907

Bradford KJ and Yang SF (1980) Xylem transport of 1-aminocyclopropane-1-carboxylic acid, an ethylene precursor, in waterlogged tomato plants. Plant Physiology 65: 322-326

Bradley DJ, Kjellbom P and Lamb CJ (1992) Elicitor- and wound-inducible oxidative cross-linking of a proline-rich plant cell wall protein. Cell 70: 21-30

Brown JWS (1986) A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Research 14, 9549-9559

Butland SL, Chow ML, Ellis BE (1998) A diverse family of phenylalanine ammonia-lyase genes expressed in pine trees and cell cultures. Plant Molecular Biology 37(1):15-24.

Cabello F, Jorrin JV and Tena M (1994) Chitinase and β-1,3-glucanase activities in chickpea (Cicer arietinum) – induction of different isoenzymes in response to wounding and ethephon. Physiologia Plantarum 92(4): 654-660

Callahan AM, Morgens PH, Wright P and Nichols KE Jr (1992) Comparision of Pch313 (pTOM13 homolog) RNA accumulation during fruit softening and wounding of two phenotypically different peach cultivars. Plant Physiology 100: 482-488

Cazzonelli CI, Cavallaro AS and Botella JR (1998) Cloning and characterization of ripening-induced ethylene biosynthetic genes from non-climacteric pineapple (Anannas comosus) fruits. Australiaan Journal of Plant Physiology 25(5): 513-518

Chavarriaga P, Mancila L, Segovia V and Roca W (1998) Developing cassava transformation strategies against stem borers. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Chen L, Zhan S, Beachy RN and Fauquet CM (1998) A protocol for consistent, large scale production of fertile transgenic rice plants. Plant Cell Report. 18(1-2): 25-31

Chen M and McClure JW (2000) Altered lignin composition in phenylalanine ammonia-lyase-inhibited radish seedlings: implications for seed-derived sinapoyl esters as lignin precursors. Phytochemistry (Oxford) 53(3): 365-370.

Christou P (1992) Genetic transformation of crop plants using microprojectile bombardment. The Plant Journal 2: 275-281

Christou P and McCabe DE (1992) Prediction of germ-line transformation events in chimeric R0 transgenic soybean plantlets using tissue-specific expression patterns. Plant Journal 2: 283-290

CIAT (1992) Cassava programme 1987-1991. Working document 116

Cock JH (1985) Cassava: new potential for a neglected crop. London, Westview Press.

Collinge DB and Slusarenko AJ (1987) Plant gene expression in response to pathogens. Plant Molecular Biology 9: 389-410

Cooke R and Cock JH (1989) Cassava crops up again. New Scientists 17: 63-68

Cooper RM (1981) Pathogen-induced changes in host ultrastructure. Plant Disease Control: Resistance and Susceptibility. (B. Solheim, J.Raa, eds.) Wiley, New York, pp. 105-142

Cooper RM, Deshappriya N, Clarckson JM and Henshaw GG (1995) Cassava bacterial blight: Pathogenicity and resistance and implications in development of resistant lines. In: Thro AM & W Roca (Eds.), The Cassava Biotechnology Network: Proceedings of the Second International Scientific Meeting. CIAT Cali, Colombia. pp. 473-484

Cornejo, Maria-Jesus; Luth, Diane; Blankenship, Kathleen M.; Anderson, Olin D.; Blechl, Ann E (1993) Activity of a maize ubiquitin promoter in transgenic rice. Plant Molecular Biology 23 (3). 567-581

Coursey DG and Booth RH (1977) Root and tuber crops. Food Crops of the Lowland Tropics. Edited by Leakley and Wills JB. Oxford Univerisity Press. 75-83

Cramer CL, Edwards K, Dron M, Liang X, Dildine SL, Bolwell GP, Dixon RA, Lamb CJ and Schuch W (1989) Phenylalanine ammonia-lyase gene organisition and structure. Plant Molecular Biology 12: 367-383

Crotwold E, Drummond B, Bowen B, Peterson T (1994) The myb-homologous P gene controls phlobene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subunit. Cell 76:543-553

Cuypers B, Schmelzer E, Hahlbrock K (1988) *In situ* localization of rapidly accumulated phenylalanine ammonia-lyase mRNA around penetration sites of *Phytophthora infestans* in potato leaves. Molecular Plant Microbe Interact 1: 157-160

Dai Z, An G (1995) Induction of nopaline synthase promoter activity by H_2O_2 has no direct correlation with salicylic acid. Plant Physiology 109(4):1191-7.

Dangl JL, Hauffe KD, Lipphardt S (1987) Parsley protoplasts retain differential responsiveness to uv-light and fungal elicitor. EMBO Journal 6: (9) 2551-2556

Data ES and Quevedo MA (1984) Pruning techniques affecting the root quality of cassava at harvest and subsequent storage. In: Tropical Root Crops: post-harvest physiology and processing. (I. Uritani, E.D. Reyes, eds.), Tokyo, pp.127-143

Delaney TP, Uknes S, Vemooij B, Friedrich L, Weymann K, Negrotto D, Gaffney T, Gut-Rella M, Kessmann H, Ward E, Ryals J (1994) A central role of salicylic acid in disease resistance. Science 266: 1247-1250

Dellaporta SL, Wood J and Hicks JB (1983) A plant DNA minipreparation: version II. Plant Molecular Biology report. 1:19-21

Diallinas G and Kanellis AK (1994) A phenylalanine ammonia-lyase gene from melon fruit-cDNA cloning, sequence and expression in response to development and wounding. Plant Molecular Biology 26(1): 473-479

Dixon RA and Lamb CJ (1990) Molecular communication in interactions between plants and microbial pathogens. Annual Review Plant Physiology and Plant Molecular Biology 41: 339-367

Dixon RA and Paiva NL (1995) Stress-induced phenylpropanoid metabolism. Plant Cell 7: 1085-1097

Dixon RA, Maria JH and Paiva NL (1995) The isoflavonoid phytoalexin pathway: from enzymes to genes to transcription factors. Physiologia Plantarum 93: 385-392

Droge-Laser W, Kaiser A, Lindsay WP, Halkier BA, Loake GJ, Doerner P, Dixon RA and Lamb C (1997) Rapid stimulation of a soybean protein-serine kinase that phosphorylates a novel bZIP DNA-binding protein, G/HBF-1, during the induction of early transcription-dependent defenses. EMBO J 17;16

Edwards K, Cramer CJ, Bolwell GP, Dixon RA, Schuch W and Lamb CJ (1985) Rapid transient induction of phenylalanine ammonia-lyase mRNA in elicitor-treated bean cells. Proceedings of the National Academy of Sciences of USA. 82: 6731-6735

Elliott KA, Shirsat AH (1998) Promoter regions of the *extA* extensin gene from *Brassica napus* control activation in response wounding and tensile stress. Plant Molecular Biology 37(4):675-87.

Esashi G (1991) Ethylene and seed germination. In. The Plant Hormone Ethylene. Eds. A.K. Mattoo and J.C. Suttle. Boca Raton, FL: CRC. PP.133-157).

Eyal Y, Meller Y, Lev-Yadum S and Fluhr R (1993) A basic-type PR-1 promoter directs ethylene responsiveness, vascular and abscission zone-specific expression. Plant Journal 4: 225-234

Fahn A (1982) Secondary Xylem, Chapter 15. In: Plant Anatomy ISBN:0080280307 311-313 Oxford press,

Fang RX, Nagy F, Sivasubramaniam S and Chua NH (1989) Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell. 1(1):141-150

Farmer EE and Ryan CA (1990) Interplant communication: Airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. Proceedings of the National Academy of Sciences of USA 87: 7713-7716

Faulkner G and Kimmins WC (1975) Staining reactions of the tissue bordering lesions induced by wounding, tobacco mosaic virus, and tobacco necrosis virus in bean. Phytopathology. 65: 1396-1399

Feldbruegge M, Sprenger M, Hahlbrock K, Weisshaar B (1997) Pcmybl, a novel plant protein containing a DNA binding domain with one Myb repeat, interacts in vivo with a light-regulatory promoter unit. Plant Journal 11: 1079-1093.

Flavell RB (1994) Inactivation of gene-expression in plants as a consequence of specific sequence duplication. Proceedings of the National Academy of Sciences of USA 91 (9): 3490-3496

Foster R, Izawa T, Chua NH (1994) Plant bZIP proteins gather at ACGT elements. FASEB Journal 8(2):192-200

Fraley, RT; Rogers,SG; Horsch, RB Sanders, PR,; Flick JS; Adams SP and Hoffman NL (1983) Expression of bacterial genes in plant cells. Proceedings of the National Academy of Sciences of USA 80, 4803-4807

Franche C; Bogusz D; Schopke C; Fauquet C and Beachy RN (1991) Transient gene expression in cassava using high-velocity microprojectiles. Plant Molecular Biology 17:493-498

Frank RL, Vodkin LO (1991) Sequence and structure of a phenylalanine ammonia-lyase gene from Glycine max. DNA Sequences 1:335-46

Fukasawa-Akada T, Kung SD, Watson JC (1996) Phenylalanine ammonia-lyase gene structure, expression, and evolution in Nicotiana. Plant Molecular Biology 30(4):711-22

Gaffney T, Friedrich L, Vernooj B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H, Ryals J (1993) Requirement of salicylic acid for the induction of systemic acquired resistance. Science 261: 754-756).

GarciaMuniz N, MartinezIzquierdo JA and Puigdomenech P (1998) Induction of mRNA accumulation corresponding to a gene encoding a cell wall hydroxyproline-rich glycoprotein by fungal elicitors. Plant Molecular Biology 38(4): 623-632

Goldsbrough AP, Albrecht H and Stratford R (1993) Salicylic acid-inducible binding of a tobacco nuclear protein to a 10 bp sequence which is highly conserved amongst stress-inducible genes. Plant Journal 3: 563-571

Gomez-Cadenas A, Tadeo FR, Primo-Millo E, Talon M (1998) Involvement of abscisic acid and ethylene in the responses of citrus seedlings to salt shock. Physiologia Plantarum 103: (4) 475-484

Gomez-Lim MA, Valdes-Lopez V, Cruz-Hernandez A and Saucedo-Arias LJ (1993) Isolation and characterization of a gene involved in ethylene biosynthesis from *Arabidopsis thaliana*. Gene 134: 217-221

Gonzalez AE, Schopke C, Taylor N, Beachy RN, Fauquet CM **(1998)** Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) through *Agrobacterium*-mediated transformation of embryogenic suspension cultures Plant Cell Reports17: 827-831

Gowri G, Paiva NL, Dixon RA (1991) Stress responses in alfalfa (Medicago sativa L.) 12 Sequence analysis phenylalanine ammonia-lyase (PAL) cDNA clones and appearance of PAL transcripts in elicitor-treated cell cultures and developing plants. Plant Molecular Biology 17(3):415-29

Gray-Mitsumune M, Molitor KE, Cukovic D, Carlson CJ and Douglas CJ (1999) Developmentally regulated patterns of expression directed by popar PAL promoters in transgenic tobacco and poplar. Plant Molecular Biology 39:657-659

Gresshoff P and Doy C (1974) Derivation of a haploid cell line from Vitis Vinifera and the importance of the stage of meiotic development of the anthers for haploid culture of this and other genera. Zeitschrift Für Pflanzenphysiologie 73: 132-141

Grierson D, Maunders MJ, Slater A, Ray J, Bird CR, Schuch W, Holdsworth ML, Tucker GA and Knapp JE 1986 Gene expression during tomato ripening. Philosophical Transactions of the Royal Society. Series B 314: 399-410

Grosskopf DG; Felix G; Boller T (1991) A yeast-derived glycopeptide elicitor and chitosan or digitonin differentially induce ethylene biosynthesis, phenylalanine ammonia-lyase and callose formation in suspension-cultured tomato cells. Journal of Plant Physiology 138: 741-746

Grotewold E, Drummond BJ, Bowen B, Peterson T (1994) The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell 76(3): 543-53

Guypers B, Schmelzer E and Hahlbrock K (1988) In situ localization of rapidly accumulated phenylalanine ammonia-lyase mRNA around penetration sites of *Phytophthora infestans* in potato leaves. Molecular Plant-Microbe interaction 1:157-160

Hafida O, Mostafa M, Marie Joseph A and Monique L (2000) Effect of gamma-irradiation on phenolic compounds and phenylalanine ammonia-lyase activity during storage in relation to peel injury from peel of Citrus clementina Hort. Ex. Tanaka. Journal of Agricultural and Food Chemistry 48 (2): 559-565.

Halbrock K and Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. Annual Review Plant Physiology and Plant Molecular Biology 40:347-369

Hahlbrock K, Scheel D, Logemann E, Nurnberger T, Parniske M, Reinold S, Sacks WR, Schmelzer E. (1995) Oligopeptide elicitor-mediated defense gene activation in cultured parsley cells. Proceedings of the National Academy of Sciences of USA 92(10):4150-7

Hamilton AJ, Bouzayen M, Grierson D. (1991) Identification of a tomato gene for the ethylene-forming enzyme by expression in yeast. Proceedings of the National Academy of Sciences of USA 88(16):7434-7.

Han Y (2000) Molecular analysis of post-harvest physiological deterioration of cassava. PhD thesis, University of Bath.

Hatton D, Sablowski R, Yung M H, Smith C, Schuch W, Bevan M (1995) Two classes of *cis* sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. Plant Journal 7:859-876

Hatton D, Smith C, Bevan M (1996) Tissue-specific expression of the PAL3 promoter requires the interaction of two conserved *cis* sequences. Plant Molecular Biology. 31(2):393-7

Hauffe KD, Paszkowsi U, Schulze-Lefert P, Hahlbrock K, Dangl L, Doulas CJ (1991) A parsley 4CL-1 promoter fragment specifies complex expression patterns in transgenic tobacco. Plant Cell 3:435-443

Hedrick SA, Bell JN, Boller T and Lamb CJ (1988) Chitinase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection. Plant Physiology 86: 182-186

Henderson SJ and Friend J (1979) Increase in PAL and lignin-like compounds as race-specific resistance responses of potato tubers to *Phytophthora infestans*. Journal of Phytopathology 94: 323-34

Hirose S (1986) Physiological studied on post-harvest deterioration of cassava roots. Japanese Agriculture Research Quarter 19(4): 241-252

Hirose S and Data ES (1984) Physiology of post-harvest deterioration of cassava roots. In: Tropical Root Crops. Post-harvest Physiology and Processing. Uritani I and Reyes ED eds, Tokyo, Japan Scientific Societies Press. pp31-51

Holdsworth MJ, Bird CR, Ray J, Schuch W and Grierson D (1987) Structure and expression of an ethylene-related mRNA from tomato. Nucleic Acid Research 15; 731-739

Howles PA, Sewalt VHJ, Paiva NL, Elkind Y, Bate NJ, Lamb C and Dixon RA (1996) Over-expression of phenylalanine ammonia-lyase in transgenic tobacco plants reveals control points for flux into phenylpropanoid biosynthesis. Plant Physiology 112: 1617-1624

Huang J, Bachem C, Jacobsen E and Visser R (1998) Molecular analysis of post-harvest deterioration in cassava. Brazilian Cassava Journal, Supplement, Fourth International Meeting Cassava Biotechnology Network, Salvador, Brazil, pp115-118

Hughes MA (1996) Plant molecular genetics. Published by: Harlow Longman (ISBN: 0582247306), vii, 236 p.

Huttly AK, Baulcombe DC (1989) a wheat alpha-amy2 promoter is regulated by gibberellin in transformed oat aleurone protoplasts. EMBO Journal 8: (7) 1907-1913.

Ichimura K and Suto K (1998) Role of ethylene in acceleration of flower senescence by filament wounding in Portulaca hybrid. Physiologia Plantarum 104(4): 603-607

Itzhaki H, Maxson JM, Woodson WR (1994) An ethylene responsive enhancer element is involved in the senescence-related expression of the carnation glutathione-S-transferase (GST1) gene. Proceedings of the National Academy of Sciences of USA 91: 8925-8929

Janas KM, Cvikrova M, Palagiewicz A, Eder J (2000) Alterations in phenylpropanoid content in soybean roots during low temperature acclimation. Plant Physiology and Biochemistry (Paris) [print] 38 (7-8) July-August, 587-593

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Molecular Biology Report 5:387-405

Jennings DL (1976) Cassava *Manihot esculenta* (Euphorbiaceae). In Evolution of Crop Plant, ed. N.W. Simmonds, 81-84

John M, Rohrig H, Schmidt J, Walden R, Schell J (1997) Cell signalling by oligosaccharides. Trends in Plant Science. 2: 111-115

Jones DA (1998) Why are so many food plants cyanogenic? Phytochemistry 47(2): 155-162

Jones DH (1984) Phenylalanine ammonia-lyase: regulation of its induction, and its role in plant development. Phytochemistry 23(7): 1349-1359

Jones H, Ooms G, Jones MG (1989) Transient gene expression in electroporated Solanum protoplasts. Plant Molecular Biology. 13(5):503-11.

Joos HJ and Hahlbrock K (1992) Phenylalanine ammonia-lyase in potato *(Solanum tuberosum* L.).Genomic complexity, structural comparison of two selected genes and modes of expression. European Journal Biochemistry 204(2):621-9

Joshi CP (1987) An inspection of the domain between putative tata box and translation start site in 79 plant genes. Nucleic Acids Research. 15 (16): 6643-6653

Kang MK, Park KS, Choi D (1998) Coordinated expression of defense-related genes by TMV infection or salicylic acid treatment in tobacco. Molecular Cells 8(4):388-92

Kawagoe Y, Campbell BR, Murai N (1994) Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein beta-phaseolin gene. Plant Journal 5(6): 885-90

Kawamata S, Shimoharai K, Imura Y, Ozaki M, Ichinose Y, Shiraishi T, Kunoh H, Yamada T (1997) Temporal and spatial pattern of expression of the pea phenylalanine ammonia-lyase gene1 promoter in transgenic tobacco. Plant Cell Physiology. 38(7):792-803.

Kawaoka A, Kawamoto T, Sekine M, Yoshida K, Takano M, Shinmyo A (1994) A *cis*-acting element and a trans-acting factor involved in the wound-induced expression of a horseradish peroxidase gene. Plant Journal 6(1): 87-97

Ke D and Saltveit ME (1989) Wound-induced ethylene production, phenolic mechanism and susceptibility to russet spotting in iceberg lettuce. Physiologia Plantarum 76: 412-418

Kim WT, Yang SF (1994) Structure and expression of cDNAs encoding 1-aminocyclopropane-1-carboxylate oxidase homologs isolated from excised mung bean hypocotyls. Planta. 194(2):223-229.

King NLR and Bradbury JH (1995) Bitterness of cassava: Identification of a new apicosyl glucoside and other compounds that affect its bitter taste. Journal of Scientific Food Agriculture 68: 223-230

Kolattukudy PE (1980) Biopolyester membranes of plants: cutin and suberin. Science 208: 990-1000

Konan NK; Sangwan RS and Sangwan-Norreel BS (1994a) Somatic embryogenesis from cultured mature cotyledons of cassava (*Manihot esculenta* Crantz) Identification of parameters influencing the frequency of embryogenesis. Plant Cell Tissue and Organ Culture 37: 91-102

Konan NK; Sangwan RS and Sangwan-Norreel BS (1994b) Efficient in vitro shoot regeneration systems in cassava (*Manihot esculenta* Crantz). Plant Breeding 113; 227-236

Koukol J and Conn EE (1961)The metabolism of aromatic compouds in higher plants:IV. Purification and properties of the phenylalanine deaminase of *Hordeum vulgare*. Journal of Biological Chemistry. 236:2692-2698

Kpemoua D; Boher B; Nicole M; Calatayud P; Geiger JP (1996) Cytochemistry of defense responses in cassava infected by *Xanthomonas campestris pv. Manihotis*. Canadian Journal of Microbiology. 42 (11): 1131-1143

Lamb C and Dixon RA (1997) The oxidative burst in plant disease resistance. Annu. Rev. Plant Physiology Plant Molecular Biology 48: 251-275

Lamb CJ (1977) Trans-cinnamic acid as a mediator of the light-stimulated increase in hydroxycinnamoyl-CoA:quinate hydroxycinnamoyl transferase. FEBS Letters 75: 37-40

Lasserre E, Bouquin T, Hernandez JA, Bull J, Pech JC, Balague C (1996) Structure and expression of three genes encoding ACC oxidase homologs from melon (*Cucumis melo* L.). Molecular & General Genetics 251(1):81-90

Lee S, Nazar RN, Powell DA and Robb J (1992) Reduced PAL gene suppression in *Verticillium*-infected resistant tomato. Plant Molecular Biology. 18: 345-352.
Legrand M and Fritig B (1993) Plant pathogenesis-related proteins and their role in defense against pathogens. Biochimie 75(8): 687-706

Legris G, Potrykus I and Puonti-Kaerlas J (1998) Engineering insect resistance in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Leja M, Mareczek A, Wojciechowska R and Rozek S (1997) Phenolic metabolism in root slices of selected carrot cultivars. ACTA Physiological Plantarum 19(3): 319-325

Leone A, Costa A, Grillo S, Tucci M, Horvath I and Vigh L (1996) Acclimation to low water potential determines changes in membrane fatty acid composition and fluidity in potato cells, Plant Cell Environment 19:1103-1109

Leon, J; Rojo, E; Sanchez-Serrano, Jose J (2001). Wound signalling in plant. Journal of Experimental Botany 52 (354). 1-9. Wound signalling in plants.

Leyva XW, Liang X, Pintor-Toro JA, Dixon RA, Lamb CJ (1992) *cis*-element combinations determine phenylalanine ammonia-lyase gene tissure-specific epression pattern Plant Cell 4:261-271

Li HQ, Huang YW, Liang CY and Guo Y (1995) Improvement of plant regeneration from cyclic somatic embryos in cassava. In Cassava Biotechnology Network: Proceedings of the 2nd International Scientific Meeting, Bogor, Indonesia, CIAT working document 150: 289-320

Li HQ, Potrykus I and Puonti-Kaerlas J (1998) Engineering leaf life length in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Li HQ, Santter C, Potrykus I and Puonti-Kaerlas J (1996) Genetic transformation of cassava (*Manihot esculenta* Crantz) Nature Biotechnology 14(6): 736-740

Li J, Ou-Lee TM, Raba R, Amundson RG and Last RL (1993) *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell 5: 171-179

Li Q, Hunt AG (1995) A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. Plant Molecular Biology 28(5): 927-34

Liang X, Abel S, Keller JA, Shen NF, Theologis A (1992) The 1-aminocyclopropane-1-carboxylate synthase gene family of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of USA 89(22):11046-50.

Liang XW, Dron M, Cramer CL, Dixon RA and Lamb CJ (1989) Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues. Journal of Biological Chemistry 264(24): 14486-14492

Liang XW, Dron M, Schmid J, Dixon RA, Lamb CJ (1989) Developmental and environmental regulation of a phenylalanine ammonia-lyase, beta-glucuronidase gene fusion in transgenic tobacco plant. Proceedings of the National Academy of Sciences of USA 86: 9284-9288

Liddle S, Hughes J and Hughes MA (1997) Analysis of a cassava root specific β-glycosidase promoter. African Journal of Root and Tuber Crops 2(1-2):158-162

Lipphardt S, Brettschneider R, Kreuzaler F, Schell J and Dangl JL (1988) UV-inducible transient expression in parsley protoplasts identifies regulatory *cis-elements* of a chimeric *Antirrhinum-majus* chalcone synthase gene. EMBO Journal 7 (13): 4027-4033.

Liu JH, LeeT, Amon SH and Reid DM (1997) Differential and wound-inducible expression of ACC oxidase genes in sunflower seedlings. Plant Molecular Biology 34(6): 923-933

Liu JJ and Pua EC (1997) Gene cloning and expression of ACC oxidase and ACC synthase in banana. Plant Physiology 114(3SS): 846

Liu ZB, Ulmasov T, Shi X, Hagen G and Guilfoyle TJ (1994) Soybean GH3 promoter contains multiple auxin-inducible elements. Plant Cell 6, 645-657

Loake GJ, Faktor O, Lamb CJ, Dixon RA. (1992) Combination of H-box [CCTACC(N)7CT] and G-box (CACCTG) *cis* elements is necessary for feedforward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate *p*-coumaric acid. Proceedings of the National Academy of Sciences U.S.A.. **89**: 9230-4.

Logemann E, Parniske M, Hahlbrock K (1995) Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proceedings of the National Academy of Sciences of USA 92(13):5905-9.

Lois R, Dietrich A, Hahlbrock K, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive *cis*-acting elements. EMBO Journal 8(6):1641-8

Lois R, Hahlbrock K (1992) Differential wound activation of members of the phenylalanine ammonia-lyase and 4-coumarate:CoA ligase gene families in various organs of parsley plants. Z Naturforsch [C] 47(1-2):90-4

Lozano JC (1979) General consideration on cassava pathology. Proceeding, Cassava Protection Workshop CIAT, Cali 1977, 17-28

Lozano LC, Cock JH and Castano J (1978) New developments in cassava storage. CIAT Series CE-14: 135-141

Lulai EC and Corsini DL (1998) Differential deposition of suberin phenolic and aliphatic domains and their roles in resistance to infection during potato tuber (*Solanum tuberosum* L.) wound-healing. Physiological and Molecular Plant Pathology 53(4): 209-222

Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA (1987) Selection of AUG initiation codons differs in plants and animals. EMBO Journal 6 (1): 43-48

MacDiarmid CW, Gardner RC (1993) A cDNA sequence from kiwifruit homologous to 1-aminocyclopropane-1-carboxylic acid oxidase. Plant Physiology 101(2):691-2.

Maher EA, Bate NJ, Ni W, Elkind Y, Dixon RA and Lamb CJ (1994) Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. Proceedings of the National Academy of Sciences of USA 91: 7802-7806

Mahungu VB; Dixon AGO and Kumbira JM (1994) Breeding cassava for multiple pest resistance in Africa. African Crop Science Journal 2:539-552

Marcotte WR, Russell SH, Quatrano RS (1989) Abscisic acid-responsive sequences from the em gene of wheat. Plant Cell. 1(10):969-76.

Marriott J, Been BO and Perkins C (1978) The aetiology of vascular discoloration in cassava roots after harvesting: association with water loss from wounds. Physiologia Plantarum 44: 38-42

Mason HS and Mullet JE (1990) Expression of 2 soybean vegetative storage protein genes during development and in response to water deficit, wounding, and jasmonic acid. Plant Cell 2(6): 569-579

Masucci JD and Schiefelbein JW (1994) The rhd6 mutation of *Arabidopsis thaliana* alters root-hair initiation through an auxin- and ethylene-associated process. Plant Physiology 106: 1335-1346

Mathews H, Schopke C, Carcamo R, Chavarriaga P, Fauquet C and Beachy RN (1993) Improvement of somatic embryogenesis and plant recovery in cassava. Plant Cell Reports 12: 328-333

Matsumoto K, Cabral GB, Teixeira JB and Rech EL (1991) Embriogenese somatica a partir de folhas imaturas de mandioca in vitro. Revista Brasileira de Fisiologia Vegetal 3: 107-110

Mauch F, Meehl JB and Staehelin LA (1992) Ethylene-induced chitinase and β-1,3-glucanase accumulate specifically in the lower epidermis and along vascular strands of bean leaves. Planta 186(3): 367-375

Mauch-Mani B, Slusaren A (1996) Production of salicylic acid precursors is a major function of phenylalanine ammonia-lyase in the resistance of Arabidopsis to *Peronospora parasitica* Plant Cell 8:203-212

MbeguiAMbeguie D, Chahine H, Gomez RM, Gouble B, Reich M, Audergon JM, Souty M, Albagnac G and FilsLycaon B (1999) Molecular cloning and expression of a cDNA encoding ACC oxidase from apricot fruit (Prunus armeniaca). Physiologia Plantarum 105(2): 294-303

McGarvey DJ, Yu H, Christoffersen RE (1990) Nucleotide sequence of a ripening-related cDNA from avocado fruit. Plant Molecular Biology. 15(1):165-7.

McGurl B, Pearce G, Orozco-Cardenas M and Ryan CA (1992) Structure, expression, and antisense inhibition of the systemin precursur gene. Science 255: 1570-1573

McMahon J, White W and Sayre RT (1995) Cyanogenic glycosides in cassava (*Manihot esculenta* Crantz). Journal of Experimental Botany 46:731-741

Meena B, Radhaieyalakshmi R, Vidhyasekaran R and Velazhahan R (2000) Effect of foliar application of Pseudomonas fluorescens on activities of phenylalanine ammonia-lyase, chitinase and beta-1,3-glucanase and accumulation of phenolics in rice. Acta Phytopathologica et Entomologica Hungarica 34 (4). 307-315.)

Meijer AH, van Dijk EL, Hoge JH.Around (1996) Novel members of a family of AT hook-containing DNA-binding proteins from rice are identified through their in vitro interaction with consensus target sites of plant and animal homeodomain proteins. Plant Molecular Biology 31(3):607-18

Memelink J, Swords KMM, Dekam RJ, Schilperoort RA, Hoge JHC and Staehelin LA (1993) Structure and regulation of tobacco extensin. Plant Journal 4(6): 1011-1022

Menkens AE, Schindler U, Cashmore AR (1995) The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends Biochemical Science 20(12): 506-10

Miki T, Yamamoto M, Nakagawa H, Ogura N, Mori H, Imaseki H, Sato T (1995) Nucleotide sequence of a cDNA for 1-aminocyclopropane-1-carboxylate synthase from melon fruits. Plant Physiology 107(1):297-8.

Minami E, Ozeki Y, Matsuoka M, Koizuka N, Tanaka Y (1998) Structure and some characterization of the gene for phenylalanine ammonia-lyase from rice plants. European Journal Biochemistry 185:19-25

Mogilner I, Orioli GA and Bletttler CM (1967) Trial to study topophysis and photoperiodism in cassava. Bonplandia. 2 265-272

Montaldo A (1973) Vascular streaking of cassava root tubers. Tropical Science 15(1): 41-46

Montgomery J, Goldman S Deikman J, Margossian L, Fischer RL (1993) Identification of an ethylene-responsive region in the promoter of a fruit ripening gene. Proceedings of the National Academy of Sciences of USA 90: 5939-5943

Morino K, Olsen OA, Shimamoto K (1999) Silencing of an aleurone-specific gene in transgenic rice is caused by a rearranged transgene. Plant Journal 17 (3): 275-285

Mukherjee A (1995) Embryogenesis and regeneration from cassava calli of anther and leaf. In Cassava Biotechnology Network: Proceedings of the 2nd International Scientific Meeting, Bogor, Indonesia, CIAT working document 150: 375-381

Mulkey TJ, Kuzmanoff KM and Evans ML (1982) Promotion of growth and hydrogen ion efflux by auxin in roots of maize pretreated with ethylene biosynthesis inhibitors. Plant Physiology 70: 186-188

Munyikwa TRI, Raemakers CCJM, Schreuder M, Kreuze J, Suurs L, Kok R, Rozeboom M, Jacobsen E and Visser RGF (1998) Introduction and expression of antisense ADPG-pyrophos-phorylase of cassava leads to decreased levels of starch and increased levels of sugars. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Munyikwa TRI, Raemakers CJJM, Schreuder M, Kok R, Schippers M, Jacobsen E, Visser RGF. (1998) Pinpointing towards improved transformation and regeneration of cassava (*Manihot esculenta* Crantz). Plant Science 135:87-101,.

Mur L, Naylor G, Warner S, Sugars J, White R, Draper J (1996) Salicylic acid potentiates defence gene expression in tissue exhibiting acquired resistance to pathogen attack. Plant Journal 9: 559-571.

Mur LA, Bi YM, Darby RM, Firek S and Draper J (1997) Compromising early salicylic acid accumulation delays the hypersensitive response and increases viral dispersal during lesion establishment in TMV-infected tobacco. Plant Journal 12(5): 1113-1126

Murakami Y, Ichinose Y, Shiraishi T and Yamada T (1997) Functional analysis of the putative *cis*-elements involved in activation by elicitor or UV and deactivation by suppressor in the promoter of a pea gene for phenylalanine ammonia-lyase (PSPAL1). Plant Cell Physiology. 38: 1403-1408.

Murashige T and Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physilogia Plantarum 15: 472-479

Nadeau JA, Zhang XS, Nair H, O'Neill SD (1993) Temporal and spatial regulation of 1-aminocyclopropane-1-carboxylate oxidase in the pollination-induced senescence of orchid flowers. Plant Physiology 103(1):31-9.

Nakashima J, Takabe K, Fujita M and Saiki H (1997) Inhibition of phenylalanine ammonia-lyase activity causes the depression of lignin accumulation of secondary wall thickening in isolated Zinnia mesophyll cell. Protoplasma 196(1-2): 99-107

Nakatsuka A, Murachi S, Okunishi H, Shiomi S, Nakano R, Kubo Y, Inaba A (1998) Differential expression and internal feedback regulation of 1-aminocyclopropane-1-carboxylate synthase, 1-aminocyclopropane-1-carboxylate oxidase, and ethylene receptor genes in tomato fruit during development and ripening. Plant Physiology 118(4):1295-305

Napoli C, Lemieux C and Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *petunia* results in reversible co-suppression of homologous gene *in trans*. Plant Cell. 2: 279-289.

Narayanaswami TC, Ramanswami NM and Sree Rangaswami SR (1995) Somatic embryogenesis and plant regeneration in cassava. In Cassava Biotechnology Network: Proceedings of the 2nd International Scientific Meeting, Bogor, Indonesia, CIAT working document 150: 324-335

Ni M, Cui DC, Gelvin SB (1996) Sequence-specific interactions of wound-inducible nuclear factors with mannopine synthase 2' promoter wound-responsive elements. Plant Molecular Biology 30: (1) 77-96

Niki, Tomoya; Mitsuhara, Ichiro; Seo, Shigemi; Ohtsubo, Norihiro; Ohashi, Yuko (1998) Antagonistic effect of salicylic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounded mature tobacco leaves. Plant and Cell Physiology 39 (5) 500-507.

Noon RA and Booth RH (1977) Nature of post-harvest deterioration of cassava roots. Transaction of British Mycology Society 69(2): 287-290

Odennell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO and Bowels DJ (1996) Ethylene as a signal mediating the wound response of tomato plants. Science 274(5294): 1914-1917

O'Donnell P; Truesdale MR; Calvert CM, Dorans A; Roberts MR; Bowles DJ (1998) A novel tomato gene that rapidly responds to wound- and pathogen-related signals. Plant Journal 14: 137-142

O'Donnell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO, Bowles DJ (1996) Ethylene as a Signal Mediating the Wound Response of Tomato Plants. Science. 274(5294):1914-7.

Ohl S, Hedrick SA, Chory J, Lamb CJ (1990) Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. Plant Cell 2:837-848

Ohme-Takagi M, Suzuki K, Shinshi H. (2000) Regulation of ethylene-induced transcription of defense genes. Plant and Cell Physiology **41**:1187-92.

Okezie BO and Kosikowski FV (1982) Cassava as a food. Critical Review Food Science Nutrition 17(3): 259-75)

Olsen K, Schall BA. (1999) The origins of cassava: A phylogeographical analysis of *Manihot esculenta*. Proceedings of the National Academy of Sciences of USA 1196:5586-5591,

Olsen KM and Schaal B (1999) Evidence on the origin of cassava: Phylogeograghy of *Manihot esculenta*. Proceedings of the National Academy of Sciences of USA 96: 5586-5591

Onwueme IC (1978) The Tropical Tuber Crops: Yam, Cassava, Sweet potato, Cocoyams. John Wiley and Sons, Chichester

Oosterhaven K, Hartmans K, Scheffer JJC and Van Der Plas LHW (1995) S-Carvone inhibits PAL activity and suberization during wound healing of potato tubers. Journal of Plant Physiology 146:288-294

Ori N, Sessa G, Lotan T, Himmelhoch S and Fluhr R (1990) A major stylar matrix polypeptide (sp41) is a member of the pathogenesis-related proteins superclass. EMBO Journal 9(11): 3429-3436

Orozco-Cardenas M and Ryan CA (1999) Hydrogen peroxidase is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway. Proceedings of the National Academy of Sciences of USA 96: 6553-6557

Osakabe R, Osakabe K, Kawai S, Katayama R, Morohoshi, N (1995) Characterization of the structure and determination of mRNA levels of the phenylalanine ammonia-lyase gene family from *Populus kitakamiensis*. Plant Molecular Biology 28, 1133-1141

Otim-Nape GW (1995) The Africa cassava mosaic virus (ACMV): a threat to food security in Africa. In cassava Biotechnology Network: Proceedings of the 2nd international scientific Meeting, Bogor, Indonesia. CIAT working document 150: 519-527

Ow DW, Jacobs JD, Howell SH (1987) Functional regions of the cauliflower mosaic virus-35s rna promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. Proceedings of the National Academy of Sciences of USA 84 (14): 4870-4874

Oyenuga VA (1968) Nigeria's Foods and Feeding Stuffs. University of Ibadan Press, Ibadan, Nigeria pp99

Pallas JA, Paiva NL, Lamb C and Dixon RA (1996) Tobacco plants epigenetically suppressed in phenylalanine ammonia-lyase expression do not develop systemic acquired resistance in response to infection by tobacco mosaic virus. The Plant Journal 10(2): 281-293

Pastugila M, Roby D, Dumas C and Cock JM (1997) Rapid induction by wounding and bacterial infection of an S gene family receptor-like kinase gene in Brassica oleracea. The Plant Cell 9:49-60

Peña-Cortês H, Sanchez-Serrano J, Rocha-Sosa M and Willmitzer L (1988) Systemic induction of proteinase inhibitor II gene expression in potato plants by wounding. Planta 174:84-89

Pearce G; Strydom D; Johnson S; Ryan CA (1991) A polypeptide from tomato leaves induces wound-inducible proteinase-inhibitor proteins. Science. 253: 895-898

Peck SC, Olson DC, Kende H (1993) A cDNA sequence encoding 1-aminocyclopropane-1-carboxylate oxidase from pea. Plant Physiology 101(2):689-90.

Pellegrini L, Rohfritsch O, Fritig B, Legrand M (1994) Phenylalanine ammonia-lyase in tobaco: Melecular cloning and gene expression during the hypersensitive reaction to tobacco mosaic virus and the response to a fungal elicitor. Plant Physiology 106: 877-886.

Pereira F; Goodwin P and Erickson L. (1999) The role of PAL gene during cassava bacterial blight and cassava bacterial necrosis. Journal of Plant Research. 112: 51-60

Petruzzelli L, Harren F, Perrone C and Reuss J (1995) On the role of ethylene in seed germination and early root growth of Pisum sativum. Journal of Plant Physiology 145: 83-86

Picton S, Gray JE and Grierson D (1995) Ethylene genes and fruit ripening. In Plant Hormones, Physiology, Biochemistry and Molecular Biology. 2nd Ed PJ Davies. Kluwer Academic Publishers. P372-394

Plumbley RA and Rickard JE (1991) Post-harvest deterioration of cassava. Tropical Science 31: 295-303

Plumbley RA, Hughes PA and Marriott J (1981) Studies on peroxidases and vascular discoloration in cassava root tissue. Journal of Scientific Food Agriculture 32:723-731

Pogson BJ, Downs CG and Davies KM (1995) Differential expression of 2 1-aminocyclopropane-1-carboxylic acid oxidase genes in Broccoli after harvest. Plant Physiology 108(2): 651-657

Potrykus I (1991) Gene transfer to plants: Assessment and perspectives. Physilogia Plantarum 79: 125-134

Puonti-Kaerlas J, Frey P and Potrykus I (1997) Development of meristem gene transfer techniques for cassava. In Proceedings of the 3nd International Scientific Meet of the Cassava Biotechnology Network, Kampala, Uganda. African Journal of Root and Tuber Crops

Purseglove JW (1968) Tropical Crops. Dicotyledons 1, Longmans Green and Co. London pp332

Raemakers C, Bessembinder J, Staritsky G, Jacobsen E and Visser R (1993) induction, germination and shoot development of somatic embryos in cassava. Plant Cell Tissue and Organ Culture 33: 151-156.

Raemakers C, Sofiari E, Taylor N, Hanshaw G, Jacobsen E and Visser RGF (1996) Production of transgenic cassava (*Manihot esculenta* Crantz) plants by particle bombardment using luciferase activity as selection marker. Molecular Breeding 2: 339-349

Rao KV, Rathore KS, Hodges TK, Fu X, Stoger E, Sudhakar D, Williams S, Christou P, Bharathi M, Bown DP, Powell KS, Spence J, Gatehouse AMR, Gatehouse JA (1998) Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. Plant Journal 15 (4): 469-477)

Rasmussen OF, Oerum H (1991) Analysis of the gene for phenylalanine ammonia-lyase from Rhodosporidium toruloides. DNA Sequences. 1(3):207-11.

Ravi V, Aked J and Balagopalan C (1996) Review on tropical root and tuber crops I. Storage methods and quality changes. Critical reviews in Food Science and Nutrition 36(7): 661-709

Reid MS (1995) Ethylene in plant growth, development, and senesence. In: Plant Hormones, Physiology, Biochemistry and Molecular Biology. 2nd edition. Ed by Davies PJ. P486-508

Reid MS, Paul JL, Farhoomand MB, Kofranek AM and Staby GL (1980) Pulse treatments with the silver thiosulfate complex extended the vase life of cut carnations. Journal of American Society of Horticulture Science 105: 25-27

Reinhardt D, Kende H and Boller T (1994) Subcellular localisation of ACC oxidase in tomato cells. Planta 195: 142-146

Reinold S, Hahlbrock K (1996) Biphasic temporal and spatial induction patterns of defense-related mRNAs and proteins in fungus-infected parsley leaves. Plant Physiology 112: 131

Rice EL (1997) Allelopathy: An overview. In Allelochemicals: Role in Agriculture and Forestry (ACS Symposium Series, no. 330), G. R. Waller, ed., american Chemical Society, Washingtom. DC., pp 8-22

Rickard JE (1981) Biochemical changes involved in the post-harvest deterioration of cassava roots. Tropical Science 23: 235-237

Rickard JE (1983) The development of occlusions in cassava root xylem vessels. Annal of Botany 52:811-821

Rickard JE (1985) Physiological deterioration of cassava roots. Journal of Sciences in Food and Agriculture 36: 167-176.

Rickard JE, Coursey DG (1981) Cassava storage. Part 1: Storage of fresh cassava roots. Tropical Science 23: 235-237

Rickard JE, Marriott J and Gahan PB (1979) Occlusions in cassava xylem associated with vascular discoloration. Annal of Botany 43: 523-526

Rickey TM and Belknap WR(1991) Comparison of the expression of several stress-responsive genes in potato tubers. Plant Molecular Biology. 16(6):1009-18

Rogers DJ and Appan SG (1973) *Manihot, Manihotoides* (*Euphorbiaceae*). Flora Neotropica (Monogragh No. 13) Hafner Press, New York

Rojo E, Leon J, Sanchez-Serrano JJ (1999) Cross-talk between wound signalling pathways determines local versus systemic gene expression in *Arabidopsis thaliana*. Plant Journal 20(2):135-142.

Rombaldi C, Lelievre JM, Latche A, Petitprez M, Bouzayen M and Pech JC (1994) Immunocytolocalization of ACC oxidase in tomato and apple fruit. Planta 192: 453-460

Ronald PR (1997) The molecular basis of disease resistance in rice. Plant Molecular Biology 35:179-186)

Ross GS, Knighton ML and Lay-Yee M (1992) An ethylene-related cDNA from ripening apples. Plant Molecular Biology 19: 231-238

Ryals J, Uknes s, Ward E (1994) Systemic acquired resistance. Plant Physiology. 104: 1109-1112

Ryder TB, Hedrick SA, Bell JN, Liang X, Clouse SD, Lamb CJ (1987) Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in phaseolus-vulgaris. Molecular & General Genetics 210(2): 219-233

Rylski I, Rappaport L and Pratt HK 1974 Dual effects of ethylene on potato dormancy and sprout growth. Plant Physiology 53: 658-662

Sablowski RW, Baulcombe DC, Bevan M (1995) Expression of a flower-specific Myb protein in leaf cells using a viral vector causes ectopic activation of a target promoter. Proceedings of the National Academy of Sciences of USA 92(15):6901-5

Sablowski RW, Moyano E, Culianez-Macia FA, Schuch W Martin C, Bevan M (1994) A flower- specific Myb protein activates transcription of phenylpropanoid biosynthetic genes EMBO Journal 13:128-137

Sambrook J, Fritsch and Maniatis T (1989) Molecular Cloning, A Laboratory Manual. 2nd edition. Cold Spring Harbour Laboratory Press

Sandberg SL and Berenbaum MR (1989) Leaf-tying by tortricid larvae as an adaptation for feeding on phototoxic hypericum perforatum. Journal of Chemical Ecology 15: 875-885

Sanders PR, Winter JA, Barnason AR, Rogers SG, Fraley RT. (1987) Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants. Nucleic Acids Research. 15(4):1543-58

Sarria, R.; Torres, E.; Angel, F.; Chavarriaga, P.; Roca, W. M (2000) Transgenic plants of cassava (*Manihot esculenta*) with resistance to Basta obtained by *Agrobacterium*-mediated transformation. Plant Cell Reports 19 (4): 339-344.

Sato F, Kitajima S, Koyama T, Yamada Y (1996) Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC *cis*-sequence. Plant Cell Physiol. 37(3):249-55.

Scharer-Hernandez NG, Frey PM, Potrykus I and Puonti-Kaerlas J (1998) Towards African cassava mosaic virus resistance in cassava. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil

Schenk R and Hildebrandt A (1972) Medium and techniques for induction and growth of monocotyledonous and dicotyledonus plant cell cultures. Canadian Journal Botany 50: 119-204

Schmelzer E, Kruger-Lebus S and Hahlbrock K (1989) Temporal and spatial patterns of gene expression around sites of attempted fungal infection in parsley leaves. Plant Cell 1: 993-1001

Schopke C, Chavarriaga P, Mathews H, Li G-G, Fauquet C and Beachy RN (1993) Transformation of cassava (*Manihot esculenta* Crantz) somatic embryos using particle bombardment, p64A, in Abstracts Congress on Cell and Tissue Cultures, San Diego, June 5-9 1993 In Vitro Cell. Development Biology. 29A

Schopke C, Taylor NJ, Carcamo R, Konan NK, Marmey P, Henshaw GG, Beachy RN and Fauguet CM (1996) Regeneration of transgenic cassava plants (*Mnihot ezculenta* Crantz) from microbombardment embrogenic suspension cultures. Nature Biotechnology 14(6): 731-735

Schopke C, Taylor NJ, Masona MV, Carcamo R, Ho T, Beachy RN and Fauguet CM (1998) Characterization of transgenic cassava plants containing the coat protein gene of cassava common mosaic virus. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Schulze-Lefert P, Dangle JL, Becker-Andre M, Hahlbrock K and Schulz W (1989) Inducible in vivo DNA foot prints define sequences necessary for UV-light activation of the parsley chalcone synthase gene. EMBO Journal. 8: 651-656

Seki S, Ichinose Y, Kato H, Shiraishi T and Yamada T (1996) Analysis of *cis*-regulatory elements involved in the activation of a member of chalcon synthase gene family (PsCHs1) in pea. Plant Molecular Biology 31:479-491

Sequin A, Laible G, Leyva A, Dixon RA and Lamb CJ (1997) Characterisation of a gene encoding a DNA-binding protein that interacts in vitro with vascular specific cis elements of the phenylalanine ammonia-lyase promoter. Plant molecular Biology 35(3): 281-91

Sessa G, Meller Y, Fluhr R (1995) A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene. Plant Molecular Biology. 28(1):145-53

Sessa G, Yang XQ, Raz V, Eyal Y, Fluhr R (1995) Dark induction and subcellular localization of the pathogenesis-related PRB-1b protein. Plant Molecular Biology. 28(3):537-47.

Shaw CH, Carter GH, Watson MD, Shaw CH (1984) A functional map of the nopaline synthase promoter. Nucleic Acids Research. 12(20):7831-46

Shiomi S, Yamamoto M, Ono T, Kakiuchi K, Nakamoto J, Natatsuka A, Kubo Y, Nakamura R, Inaba A, Imaseki H (1998) cDNA cloning of ACC synthase and ACC oxidase genes in cucumber fruit and their differential expression by wounding and auxin. Journal of the Japanese Society for Horticultural Science 67(5): 685-692

Shirasu K, Nakajima H, Rajasekhar VK, Dixon RA and Lamb C (1997) Salicylic acid potentiates an agonist-dependent gain control that amplifies pathogen signals in the activation of defense mechanisms. Plant Cell 9(2): 261-270

Showalter AM; Bell JN; Cramer CL; Bailey JA; Barner JE; Lamb JC (1985) Accumulation of hydroxyproline-rich glycoprotein mRNAs in response to fungal elicitor and infection. Proceedings of National Academy of Science of United States of America 82: 6551-6555

Shufflebotton D, Edwards K, Schuch W, Bevan M (1993) Transcription of two members of a gene family encoding phenylalanine ammonia-lyase leads to remarkably different cell specificities and induction pattern. Plant Journal 3:835-845

Siefert F, Langebartels C, Boller T and Grossmann K (1994) Are ethylene and ACC involved in the induction of chitinase and β-1,3-glucanase activity in sunflower cell-suspension cultures. Planta 192(3): 431-440

Slater A, Maunders MJ, Edwards K, Schuch W and Grierson D (1985) Isolation and characterisation of cDNA clones for tomato polygalacturonase and other ripening-related proteins. Plant Molecular Biology 5: 137-147

Smart CM. (1994) Tansley review no. 64 gene expression during leaf senescence. New Phytologist 126 (3): 419-448.

Smith CJ, Watson CF, Bird CR, Ray J, Schuch W, Grierson D. (1990) Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. Molecular & General Genetics. 224(3):477-81.

Smith CJS, Slater A and Grierson D (1986) Rapid appearance of an mRNA correlated with ethylene synthesis encoding a protein of molecular weight 35,000. Planta 168: 94-100

Smith CJS, Watson CF, Bird CR, Ray J, Schuch W and Grierson D (1990) Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. Molecular & General Genetics. 224: 477-481.

Smith DA (1982) Toxicity of phytoalexins. In Phytoalexins, JA.Bailey and Jw mansfied, eds (New York: John Wiley and Sons), pp. 218-252

Smith-Becker J, Marois E, Huguet EJ, Midland SL, Sims JJ and Keen NT (1998) Accumulation of salicylic acid and 4-hydroxybenzoic acid in phloem fluids of cucumber during systemic acquired resistance is preceded by a transient increase in PAL activity in petioles and stems. Plant Physiology (Rockville) 116 (1) : 231-238.

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C. & Ronald, P. (1995) A Receptor Kinase-Like Protein Encoded by the Rice Disease Resistance Gene, *Xa21*. Science 270, 1804-1806

Splittstoesser WE. (1992) Crop physiology of cassava. Horticulture Reviews 13:105-129,

Stadnik MJ and Buchenauer H (2000) Inhibition of phenylalanine ammonia-lyase suppresses the resistance induced by benzothiadiazole in wheat to Blumeria graminis f. sp. tritici. Physiological and Molecular Plant Pathology 57 (1): 25-34

Stamp FA and Henshaw GG (1982) Somatic embryogenesis in cassava. Z Planzenphysiologie 105:183-187

Stamp FA and Henshaw GG (1987) Somatic embryogenesis from clonal leaf tissue of cassava. Annals of Botany 53: 445-450

Stead AD (1992) Pollination induced flower senescence: a review. Plant Growth Regulation. 11: 13-20

Suarez MC, Gutierrez J, Thome J, Fregene M (2000) Developing expressed sequence tags (ESTs) from polymorphic transcript-derived fragments (TDFs) in cassava (*Manihot esculenta* Crantz). Genome. 43 (1): 62-67

Subramanian R, Reinold S, Molitor Ek and Douglas CJ (1993) Structure, inheritance, and expression of hybrid poplar (*Populus tricocarpa x Populus deltoides*) phenylalanine ammonia-lyase genes. Plant Physiology 102: 71-83

Suhandono S, Hughes J, Brown K, Sirju-Charan G and Hughes MA (1998) Characterization of an elongation factor-la gene from cassava: a new promoter for biotechnology. The Fourth International Scientific Meeting Cassava Biotechnology Network, Brazil.

Szabados L, Hoyos R and Roca WM (1987) *In vitro* somatic embryogenesis and plant regeneration of cassava. Plant Cell Reports 6:248-251

Taiz L and Zeiger E (1998) plant physiology (Second Edition). Sinauer Associates, Inc., Publishers. Sunderland, Massachusetts p349-376

Takeda J, Ozeki Y and Yoshida K (1997) Action spectrum for induction of promoter activity of phenylalanine ammonia-lyase gene by UV in carrot suspension cells. Photochemistry & Photobiology 66(4): 464-470

Tanaka Y, Data ES, Hirose S, Taniguchi T and Uritani I (1983) Biochemical changes in secondary metabolites in wounded and deteriorated cassava roots. Agricultural Biology and Chemistry 47(4): 693-700

Tanaka Y, Data ES, Lape VG, Villegas-Godoy CDV, Gorgonio MA and Hirose S (1984) Effect of pruning treatment on physiological deterioration and related biological changes in cassava roots. In Tropical Root Crops: Postharvest Physiology and Processing pp119-125

Tanaka Y, Matsuoka M, Yamanoto N, Ohashi Y, Kano-Murakami Y, Ozeki Y (1989) Structure and characterization of a complementary DNA clone for phenylalanine ammonia-lyase from cut-injured roots of sweet potato. Plant Physiology (Bethesda) 90(4): 1403-1407

Tang X, Wang H, Brandt AS, Woodson WR (1993) Organization and structure of the 1-aminocyclopropane-1-carboxylate oxidase gene family from Petunia hybrida. Plant Molecular Biology. 23(6):1151-64.

Taylor NJ, Edwards M, Kiernan R, Davey CDM, Blakesley D and Henshaw GG (1996) Development of friable embryogenic callus and embryogenic suspension culture systems in cassava (*Manihot esculenta* Crantz) Nature Biotechnology 14: 126-130

Taylor NJ, Kiernan RJ, Blakesley D, Henshaw GG (1997) Strategies for the production of friable embryogenic tissues across a range of cassava cultivars. In Proceedings of the 3rd International Scientific Meet of the Cassava Biotechnology Network, Kampala, Uganda. African Journal of Root and Tuber Crops

Taylor NJ, Schöpke C, Masona MV, Carcamo R, Ho Thao, Beachy RN, Fauquet CM. (1998) Production of transgenic cassava plants with marker genes and genes of interest by microparticle bombardment and *Agrobacterium*-mediated gene transfer. Brazilian Cassava Journal, Supplement, Fourth International Meeting Cassava Biotechnology Network, Salvador, Brazil, p34.

Taylor NJ, Masona MV, C Shopke and Fanquet CM (1999) Transgenic cassava for food security and economic development. In Transgenic Foods. In Press

Theologis A (1992) One rotten apple spoils the whole bushel: the role of ethylene in fruit ripening. Cell 70: 181-184

Titarenko E, Rojo E, Leon J, Sanchez-Serrano JJ (1997) Jasmonic acid-dependent and -independent signaling pathways control wound-induced gene activation in *Arabidopsis thaliana*. Plant Physiology 115(2):817-26.

TomasBarberan FA, LoaizaVelarde J, Bonfanti A and Saltveit ME (1998) Early wound- and ethylene-induced changes in phenylpropanoid metabolism in harvested lettuce. Journal of the American Society for Horticultural Science 122(3): 399-404

Uritani I, Data ES and Tanaka Y (1984) Biochemistry of postharvest deterioration of cassava and sweet potato roots. In Uritani I & Reyes ED (Eds), Tropical Root Crops: Postharvest Physiology and Proceeding, Japanese Scientific Societies Press, Tokyo. Pp 61-75

Uritani I, Data ES, Villegas RJ, Flores P and Hirose S (1983) Relationship between secondary metabolism changes in cassava root tissue and physiological deterioration. Agricultural Biology and Chemistry 47(7): 1591-1598

Van der Krol AR, Mur LA, Beld M, Mol JN and Stuitje AR (1990) Flavonoid genes in *Petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell 2: 291-9.

Vanderhee MD, Lemmers R and Bol JF (1993) Analysis of regulatory elements involved in stress-induced and organ-specific expression of tobacco acidic and basic β-1,3-glucanase genes. Plant Molecular Biology 21(3): 451-461

Vasil I (1995) Cellular and molecular improvement of cereals. Pp5-18 in Current Issues in Plant Molecular and Cellular Biology. Proceedings of the VIIIth International Congress on Plant Tissue and Cell Culture, Italy, 12-17 June 1994 Terzi, M *et al* (eds.). Kluwer Adacemic Publishers, Dordrecht, The Netherlands.

Verdaguer B, de Kochko A, Fux CI, Beachy RN and Fauquet C (1998) Functional organization of the cassava vein mosaic virus (CsVMV) promoter. Plant Molecular Biology 37: 1055-1067

Vogelilange R, Frundt C, Hart CM, Nagy F and Meins F (1994) Development, hormonal, and pathogenesis-related regulation of the tobacco class-I β-1,3-glucanase-B promoter. Plant Molecular Biology 25(2): 299-311

Wang H, Woodson WR (1991) A flower senescence-related messenger RNA from carnation shares sequence similarity with fruit ripening-related messenger RNAS involved in ethylene biosynthesis. Plant Physiology (Bethesda) 96(3): 1000-1001

Wang TW, Arteca JM, Arteca RN (1994) A 1-aminocyclopropane-1-carboxylate oxidase cDNA sequence from Pelargonium. Plant Physiology 106(2):797-8.

Wanner LA, Li G, Ware D, Somssich IE and Davis KR (1995) The phenyalanine ammonia-lyase gene family in *Arabidopsis thaliana*. Plant Molecular Biology 27: 327-338

Watson JD, Gilman M, Witkowski J and Zoller M (1992) Recombinant DNA (second edition) Scientific American Books.

Wen CM, Wu M, Goh CJ, Pua EC (1993) Nucleotide sequence of a cDNA clone encoding 1-aminocyclopropane-1-carboxylate synthase in mustard (Brassica juncea [L.] Czern & Coss). Plant Physiology 103(3):1019-20.

Werr W, Lorz H 1986 Transient gene-expression in a gramineae cell-line - A rapid procedure for studying plant promoters Molecular & General Genetics 202 (3): 471-475

Wheatley C, Fernandez FO and Medina LML (1989) Conservation of cassava roots in polythene bags. CIAT, Cali Colombia pp33

Wheatley CC and Schwabe WW (1985) Scopoletin involvement in post-harvest physiological deterioration of cassava root (*Manihot esculenta* Crantz). Journal of Experimental Botany 36: 783-791

Whetten RW and Sederoff RR (1992) Phenylalanine ammonia-lyase from loblolly pine: purification of the enzyme and isolation of complementary DNA clones. Plant Physiology 98: 380-386

White WLB, Arias-Garzon DI, McMahon JM and Sayre RT (1998) Cyanogenesis in cassava-The role of hydroxynitrile lyase in root cyanide production. Plant Physiology 116: 1219-1225

Wu SC, Hahlbrock K (1992) In situ localization of phenylpropanoid-related gene expression in different tissues of light-grown and dark-grown parsley seedlings. Z Naturforsch C 47: 591-600

Yamada T, Tanaka Y, Sriprasertsak P, Kato H, Hashimoto T, Kawamata S, Ichinose Y, Kato H, Shiraishi T, Oku H (1992) phenylalanine ammonia-lyase genes from *Pisum sativum*: structure, organ-specific expression and regulation by fungal elicitor and suppressor. Plant cell physiology 33: 715-725

Yang SF (1987) The role of ethylene and ethylene synthesis in fruit ripening. In Plant senescence: its biochemistry and physiology. Eds. WW Thompson, EA Nothnagel and RC Huffaker. American Society of Plant Physiologists. P 156-166

Yang SF, Hoffman NE (1984) Ethylene biosynthesis and its regulation in higher-plants. Annual Review Plant Physiology 35: 155-189

Yeoh, HH; Truong, VD (1996) Protein contents, amino acid compositions and nitrogen-to-protein conversion factors for cassava roots. Journal of the Science of Food and Agriculture 70 (1). 51-54

Yu LM, Lamb CJ, Dixon RA (1993) Purification and biochemical characterization of proteins which bind to the H-box *cis*-element implicated in transcriptional activation of plant defense genes. Plant Journal 3(6):805-16

Zacarias L and Reid MS (1990) Role of growth regulators in the senescence of Arabisopsis thaliana leaves. Physiologia Plantarum: 80: 549-557

Zarembinski TI and Theologis A (1994) Ethylene biosynthesis and action: a case for conservation. Plant Molecular Biology 26: 1579-1597

Zhang S, Song WS, Chen L, Ruan D, Taylor N, Ronald P, Beachy R and Fauquet C (1998) Transgenic elite Indica rice varieties , resistant to *Xanthomonas oryzae pv oryzae*. Molecular Breeding 4(6): 551-558.

Zhang P, Legris G, Coulin P, Puonti-Kaerlas J (2000) Production of stably transformed cassava plants via particle bombardment. Plant Cell Reports. 19 (10): 939-945

Zhang W; Mcelroy D; Wu R (1991) :Analysis of rice *act1* 5' region activity in transgenic rice plant. Plant Cell 3 (11) 1155-1166.

Zhu Q, Dabi T, Beeche A, Yamamoto R, Lawton MA and Lamb C (1995) Cloning and properties of a rice gene encoding phenylalanine ammonia-lyase. Plant Molecular Biology 29: 535-550